Schoenborn, B. P. (1965) *Nature (London) 208*, 760–762.
Schoenborn, B. P. (1967) *Nature (London) 214*, 1120–1122.
Schoenborn, B. P., Watson, H. C., & Kendrew, J. C. (1965) *Nature (London) 207*, 28–30.
Smith, J. L., Hendrickson, W. A., Honzatko, R. B., & Sheriff, S. (1986) *Biochemistry 25*, 5018–5027.
Ten Eyck, L. F. (1977) *Acta Crystallogr., Sect. A: Cryst. Phys., Diffr., Theor. Gen. Crystallogr. A33*, 486–492.
Tilton, R. F. (1988) *J. Appl. Crystallogr. 21*, 4–9.
Tilton, R. F., Jr., Kuntz, I. D., Jr., & Petsko, G. A. (1984) *Biochemistry 23*, 2849–2857.
Tilton, R. F., Jr., Singh, U. C., Weiner, S. J., Connolly, M. L., Kuntz, I. D., Jr., Kollman, P. A., Max, N., & Case, D. A. (1986) *J. Mol. Biol. 192*, 443–456.
Tilton, R. F., Singh, U. C., Kuntz, I. D., & Kollman, P. A. (1988) *J. Mol. Biol. 199*, 195–211.
Wagner, G. (1980) *FEBS Lett. 112*, 280–284.
Wishnia, A. (1969) *Biochemistry 8*, 5064–5070.
Woodward, C. K., & Hilton, B. D. (1979) *Annu. Rev. Biophys. Bioeng. 8*, 99–127.
Zipp, A., & Kauzmann, W. (1973) *Biochemistry 12*, 4217–4228.

Structural Comparison of Two Serine Proteinase–Protein Inhibitor Complexes: Eglin-C–Subtilisin Carlsberg and CI-2–Subtilisin Novo†

Catherine A. McPhalen‡ and Michael N. G. James*

Medical Research Council of Canada Group in Protein Structure and Function, Department of Biochemistry, University of Alberta, Edmonton, Alberta, Canada T6G 2H7

Received November 10, 1987; Revised Manuscript Received March 24, 1988

ABSTRACT: The crystal structures of the molecular complexes between two serine proteinases and two of their protein inhibitors have been determined: subtilisin Carlsberg with the recombinant form of eglin-c from the leech *Hirudo medicinalis* and subtilisin Novo with chymotrypsin inhibitor 2 from barley seeds. The structures have been fully refined by restrained-parameter least-squares methods to crystallographic R factors ($\sum||F_o| - |F_c||/\sum|F_o|$) of 0.136 at 1.8-Å resolution and 0.154 at 2.1-Å resolution, respectively. The 274 equivalent α-carbon atoms of the enzymes superpose with an rms deviation of 0.53 Å. Sequence changes between the enzymes result in localized structural adjustments. Functional groups in the active sites superpose with an rms deviation of 0.19 Å for 161 equivalent atoms; this close similarity in the conformation of active-site residues provides no obvious reason for known differences in catalytic activity between Carlsberg and Novo. Conformational changes in the active-site region indicate a small induced fit of enzyme and inhibitor. Some conformational differences are observed between equivalent active-site residues of subtilisin Carlsberg and α-chymotrypsin. Despite differences in tertiary architecture, most enzyme–substrate (inhibitor) interactions are maintained. Subtilisin Carlsberg has a rare *cis*-peptide bond preceding Thr211 (Gly211 in Novo). Both enzymes contain tightly bound Ca^{2+} ions. Site 1 is heptacoordinate with the oxygen atoms at the vertices of a pentagonal bipyramid. Site 2 in Carlsberg is probably occupied by a K^+ ion in Novo. Conserved water molecules appear to play important structural roles in the enzyme interior, in the inhibitor β-sheet, and at the enzyme–inhibitor interface. The 62 equivalent α-carbon atoms of the inhibitors superpose with an rms deviation of 1.68 Å. Sequence changes result in somewhat different packing of the α-helix, β-sheet, and reactive-site loop relative to each other. Hydrogen bonds and electrostatic interactions supporting the conformation of the reactive-site loop are conserved. The 24 main-chain plus C^β atoms of P_4 to P_1' overlap with an rms deviation of 0.19 Å. Features contributing to the inhibitory nature of eglin-c and CI-2 are discussed.

Subtilisin Carlsberg was the first bacterial serine proteinase to be discovered (Linderstrøm-Lang & Ottesen, 1947). It consists of a single polypeptide chain of 274 amino acids, M_r 27 292 (Smith et al., 1968) with extensive sequence identity with subtilisn Novo (Figure 1a). Recently, solution of the crystal structure of proteinase K (Pähler et al., 1984) demonstrated that this fungal enzyme also belongs to the subtilisin family. Subtilisins Novo and Carlsberg have broad substrate specificities with preferences for large aromatic or aliphatic amino acids in the P_1 position of the substrate.[1] The catalytic triad, histidine, aspartic acid, and serine, has a similar but not identical spatial arrangement in the subtilisins and the chymotrypsin family of serine proteinases (Kraut et al., 1972; Wright, 1972). These two families of enzymes form one of the best-known examples of convergent evolution, with completely different tertiary structures providing the framework for a similar disposition of active-site residues.

Eglin-c is a protein inhibitor of serine proteinases isolated from the leech *Hirudo medicinalis* (70 amino acids, M_r 8092;

† This research was generously supported by a group grant from the Medical Research Council of Canada. C.A.M. gratefully acknowledges the Alberta Heritage Foundation for Medical Research for financial support. The work described in this paper forms part of the doctoral thesis of C.A.M.

‡ Present address: Department of Structural Biology, Biozentrum, University of Basel, CH-4056 Basel, Switzerland.

[1] Nomenclature of Schechter and Berger (1967). Amino acid residues of substrates or inhibitors are numbered P_1, P_2, P_3, etc. toward the N-terminus and P_1', P_2', etc. toward the C-terminus from the reactive-site bond (the peptide bond in the substrate or inhibitor cleaved by the enzyme). The complementary subsites of the binding region on the enzyme are numbered S_1, S_2, etc. and S_1', S_2', etc.

0006-2960/88/0427-6582$01.50/0

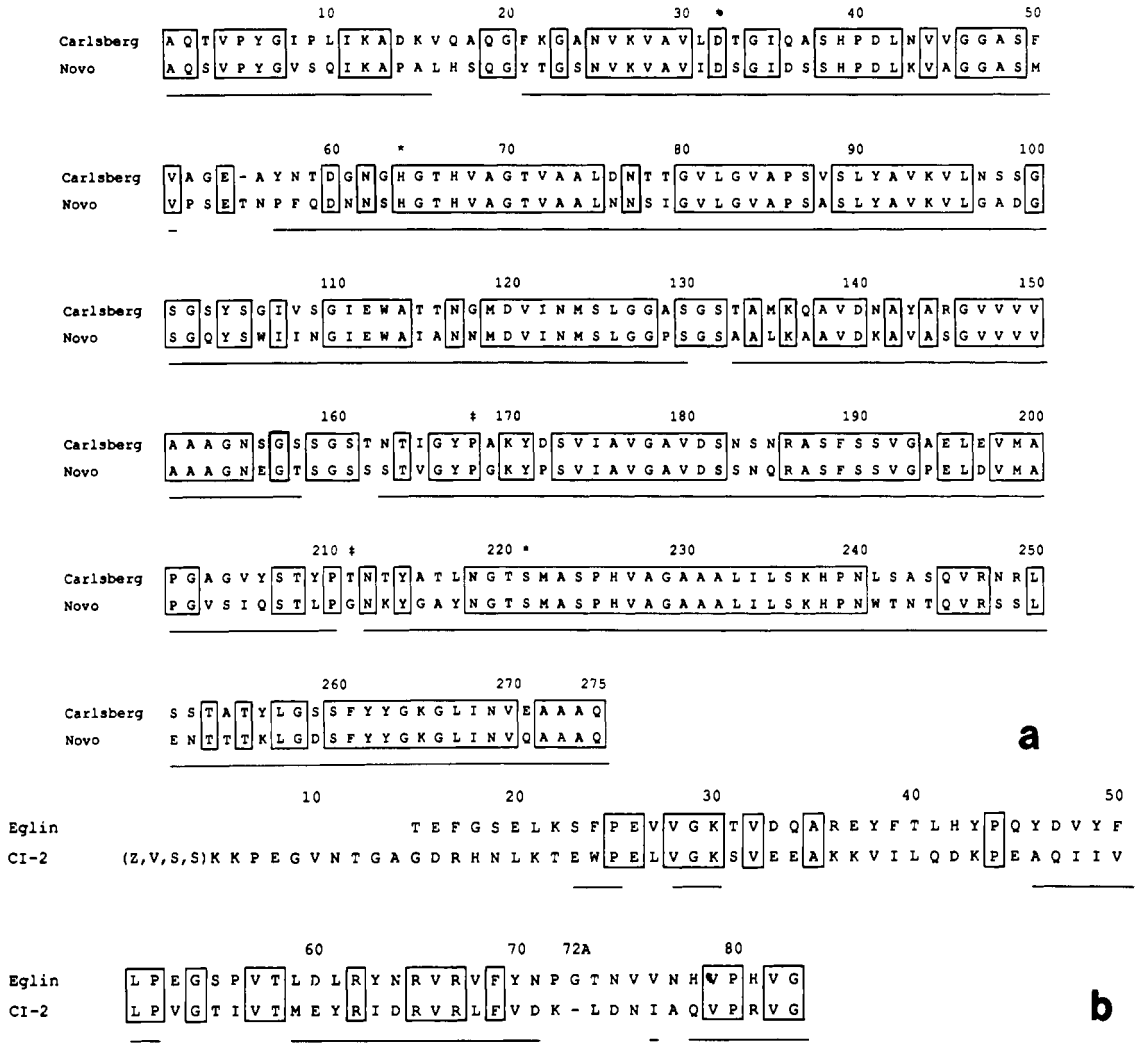

FIGURE 1: (a) Amino acid sequences of subtilisins Novo and Carlsberg. Subtilisin Novo is chemically identical with subtilisin BPN′. The amino acid sequence of subtilisin Novo, deduced from the DNA sequence (Wells et al., 1983), shows the following changes from the published protein sequence (Olaitan et al., 1968): Pro56 → Asn56, Asn57 → Pro57, Asp61 → Asn61, Ser88 → Ala88, Ala89 → Ser89, Asp98 → Ala98, Ala99 → Asp99, Ser158 → Thr158, Thr159 → Ser159, Gln251 → Glu251. The sequence of subtilisin Carlsberg shows one change from the published protein sequence (Smith et al., 1968): Asn158 → Ser158. These changes are observed in the electron density maps of the two enzyme structures. There are 82 amino acid changes between the two enzymes and one deletion; thus, they are 70% identical. Spatial equivalence places the deletion at Thr55 (Novo), although it was assigned to Pro56 (Novo) on the basis of sequence homology alone. Identical residues are boxed. Residues with rms deviations in α-carbon position of less than 1.0 Å after superposition of the molecules are underlined. Residues of the catalytic triad are denoted with an asterisk. *cis*-Peptide bonds (‡) occur between Tyr167 and Pro168 in both enzymes and between Pro210 and Thr211 in Carlsberg. (b) Amino acid sequences of CI-2 and eglin-c. The amino acid sequence of CI-2 was determined by Svendsen et al. (1980) and that of eglin-c by Seemüller et al. (1980). CI-2 is subject to hydrolysis during purification between Asn11I and Thr12I, Gly15I and Asp16I, and Arg17I and His18I (Svendsen et al., 1980). The inhibitors have 31% sequence identity.

Seemüller et al., 1977, 1980) (Figure 1b). It is a member of the potato inhibitor 1 (PI-1)[2] family of inhibitors, most of whose members lack stabilizing disulfide bonds (Melville & Ryan, 1972; Svendsen et al., 1982). Eglin-c is a strong inhibitor of mammalian leukocyte elastase, cathepsin G, and chymotrypsin ($K_i \simeq 10^{-11}$ M) but not of other mammalian proteinases tested (Schnebli et al., 1985). It forms very tight inhibitory complexes with several subtilisins. On the basis of its inhibitory properties, it is a potential candidate for the treatment of disease conditions such as emphysema and septicemia (Seemüller et al., 1986). In the kinetic studies and the present structure determination, the genetically engineered product *N*-acetyleglin-c was used (Rink et al., 1984). Chymotrypsin inhibitor 2 (CI-2) from the seeds of the Hiproly strain of barley is a second member of the PI-1 family (Svendsen et al., 1982). CI-2 is 31% identical in sequence with eglin-c (Figure 1b) and also lacks disulfide bonds. It is a strong inhibitor of subtilisins and chymotrypsins but not of trypsin (Svendsen et al., 1982).

The molecular structures of eglin-c and CI-2 have been solved in complex with two different subtilisins: eglin-c with subtilisin Carlsberg (EC–SC) (McPhalen et al., 1985a) and CI-2 with subtilisin Novo (CI-2–SN) (McPhalen et al., 1985b). Both complexes have been further refined subsequent to the initial structure reports (McPhalen, 1986). The final crystallographic *R* factors[3] are 0.136 for EC–SC (data in the

[2] Abbreviations: α_c, calculated structure factor phase; *B*-factor, thermal motion parameter = $8\pi^2\bar{U}^2$, where $\bar{U}^2$ is the mean square amplitude of vibration; CI-2, chymotrypsin inhibitor 2 from barley seeds; CI-2–SN, molecular complex between inhibitor CI-2 and subtilisin Novo; EC–SC, molecular complex between inhibitor eglin-c and subtilisin Carlsberg; $|F_o|$, $|F_c|$, observed and calculated structure factor amplitudes; I, an I follows the residue numbers of amino acids of the inhibitors to distinguish them from those of the enzymes; OMTKY3, third domain of the turkey ovomucoid inhibitor; PI-1, potato inhibitor 1; rms, root mean square; SGPA, SGPB, *Streptomyces griseus* proteinases A and B; SSI, *Streptomyces* subtilisin inhibitor.

[3] $R = \sum||F_o| - |F_c||/\sum|F_o|$, where $|F_o|$ and $|F_c|$ are the observed and calculated structure factor amplitudes, respectively.

resolution range 8.0–1.8 Å) and 0.154 for CI-2–SN (8.0–2.1 Å). The crystal structure of uncomplexed native CI-2 has been completed at 2.0-Å resolution (McPhalen & James, 1987). Comparison of the free CI-2 structure with that of its complexed form reveals increased ordering and some conformational changes of residues near the reactive site that take place on forming complexes with subtilisins. A second determination of the structure of an EC–SC complex at high resolution has been published recently (Bode et al., 1986, 1987).

Crystallographic (Read & James, 1986) and biochemical studies of other families of serine proteinase inhibitors show that they perform their function by means of tight binding to, and slow release from, their cognate enzymes (Laskowski & Kato, 1980). Eglin-c and CI-2 are believed to function via the same mechanism, but without the stabilizing disulfide bridges that flank the reactive-site region in other inhibitor families. Despite high sequence identities among reactive-site regions of different inhibitor families, regions distant from the reactive site show little sequence or structural similarity.

In this paper we present analyses of the fully refined three-dimensional X-ray crystallographic structures of the EC–SC and CI-2–SN complexes. The structures of the two enzymes and inhibitors are compared, as well as the two sets of enzyme–inhibitor interactions. Factors contributing to the inhibitory nature of eglin-c and CI-2 are discussed. The refined coordinates of both complexes have been deposited with the Brookhaven Protein Data Bank (Bernstein et al., 1977).

Methods and Results

Crystals of the complexes EC–SC and CI-2–SN were grown as described previously (McPhalen et al., 1985a,b). X-ray intensity data collection and processing were carried out as for free CI-2 (McPhalen & James, 1987). The maximum corrections for absorption and decay respectively were for EC–SC 1.48 and 1.42 and for CI-2–SN 2.09 and 1.11. The method of molecular replacement (Rossmann, 1973) was used to solve the phase problem for both complexes (McPhalen et al., 1985a,b), using the fast rotation function (Crowther, 1973; program modified by E. Dodson) and the translation function of Fujinaga and Read (1987). All electron density maps calculated for these structures had coefficients $2m|F_o| - D|F_c|$, phases α_c, where m and D are factors to suppress model bias resulting from phasing by partial structures with errors (Read, 1986). The MMS-X interactive graphics (Barry et al., 1976) with the macromolecular modeling system M3 (Sielecki et al., 1982) was used for map interpretation and model fitting.

The restrained-parameter least-squares refinement method of Hendrickson and Konnert (1980) was used to refine the atomic parameters of EC–SC and CI-2–SN. The parameters restrained and the general strategy of refinement are similar to those for penicillopepsin (James & Sielecki, 1983) and SGPA (Sielecki et al., 1979). The refinement was carried out on the FPS164 attached processor, using a version of the refinement program PROLSQ modified locally by M. Fujinaga.

The refinement parameters from the final cycle for EC–SC and CI-2–SN are given in Table I. The restraints on the final structures are tight, and the rms deviations from ideal values are small. Difference electron density maps with coefficients $m|F_o| - |F_c|$, phases α_c, were computed at the completion of refinement for both complexes and were essentially featureless. The accuracy of the atomic coordinates in the fully refined EC–SC and CI-2–SN structures was estimated with the σ_A plot of Read (1986). The σ_A plot gives an overall estimate of the coordinate errors in the structure based on a derivation similar to that of Luzzati (1952), but using more realistic basic assumptions. The overall mean coordinate error for EC–SC from this method is 0.07 Å; that for CI-2–SN is 0.21 Å. Figure 2 shows mean residue B-factors as a function of position along the polypeptide chain for each complex. It may be used to obtain an estimate of the accuracy of a specific region for each structure. Regions with main-chain B-factors above the mean for a complex are all external loops. Residues forming the active sites of the enzymes and the reactive-site loops of the inhibitors have some of the lowest B-factors in the complexes.

Secondary structural elements of the proteins were defined in two ways: first, on the basis of characteristic hydrogen-bonding patterns (maximum donor–acceptor distance of 3.4 Å, maximum donor–proton–acceptor angle of 125°) and ϕ, ψ angles; second, by the electrostatic interaction energy method of Kabsch and Sander (1983). The agreement between these two methods for both complex structures is very good.

The main-chain torsion angles ϕ and ψ for most residues in both complexes fall well within conformationally allowed

Table I: Final Refinement Parameters and Results

	EC–SC	CI-2–SN
no. of cycles	57	71
R factor	0.136	0.154
resolution range (Å)	8.0–1.8	8.0–2.1
no. of reflections [$I > \sigma(I)$]	27094	16128
no. of protein atoms	2450	2451
no. of solvent atoms[a]	167	167
no. of variable parameters	10636	10640
$\langle \|\|F_o\| - \|F_c\|\| \rangle$	17	67
⟨coordinate shift⟩ (Å) in final cycle	0.005	0.012
⟨B factor shift⟩ (Å^2) in final cycle	0.22	0.25
rms deviations from ideal values[b]		
distance restraints (Å)		
bond distance	0.009 (0.008)	0.006 (0.008)
angle distance	0.027 (0.016)	0.024 (0.016)
planar 1–4 distance	0.031 (0.016)	0.018 (0.016)
plane restraint (Å)	0.017 (0.012)	0.013 (0.012)
chiral-center restraint (Å^3)	0.146 (0.080)	0.119 (0.080)

[a] Including Ca^{2+} ions in the total. [b] The values of σ, in parentheses, are the input estimated standard deviations that determine the relative weights of the corresponding restraints [see Hendrickson and Konnert (1980)].

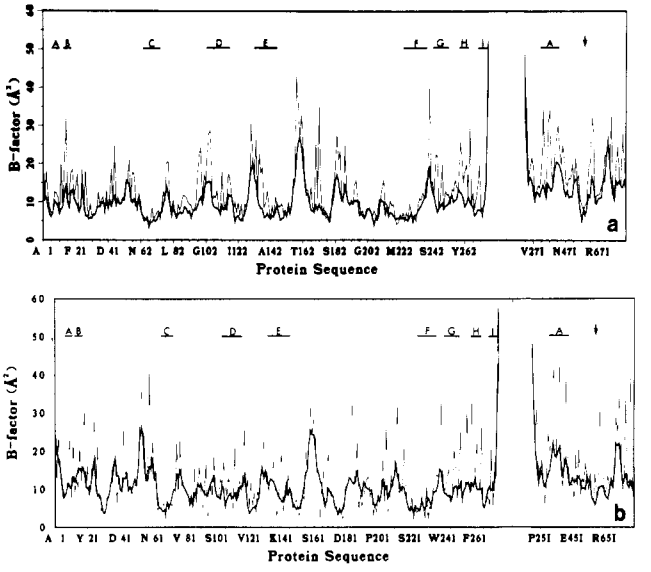

FIGURE 2: Variation in B factor along the polypeptide chain. Heavy lines denote the mean residue B factors of main-chain atoms and light lines those of the side-chain atoms. Vertical lines separate enzyme and inhibitor residues. No B factors are given for inhibitor residues before 21I; they are not seen in either electron density map. The mean overall B factor for the atoms of EC–SC is 12 Å^2 and that for CI-2–SN is 13 Å^2. Horizontal bars indicate α-helical regions; the arrows mark the reactive sites of the inhibitors. (a) EC–SC. (b) CI-2–SN.

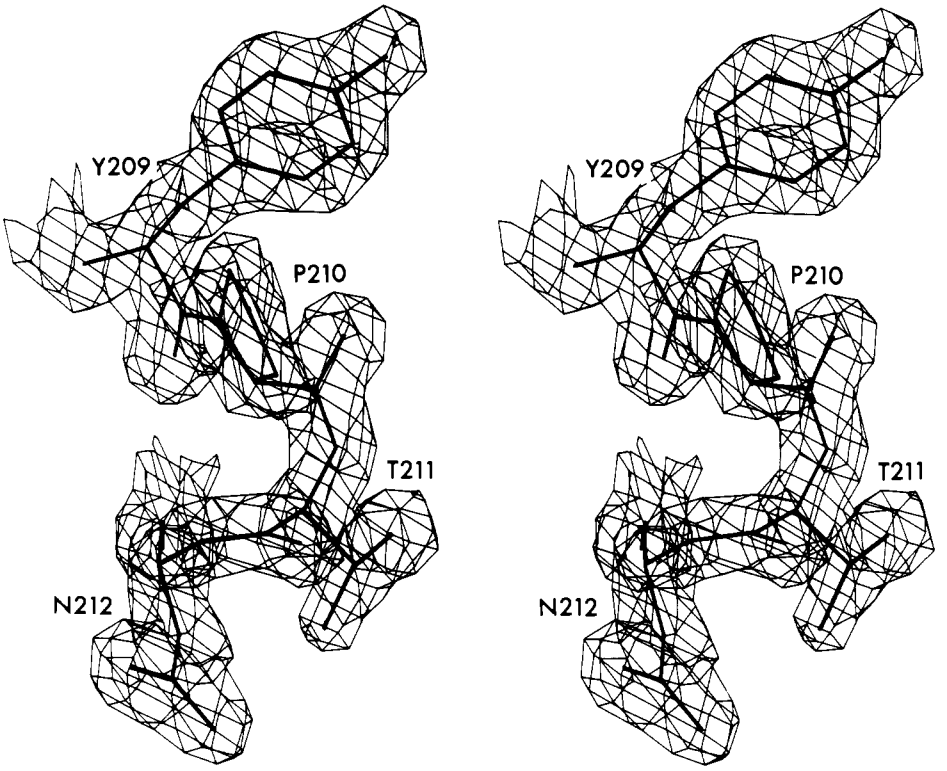

FIGURE 3: *cis*-Peptide bond preceding Thr211. Stereoview of the electron density in the region of the second *cis*-peptide bond in subtilisin Carlsberg between Pro210 and Thr211. The electron density was contoured at a level of 0.60 e/Å^2. The corresponding residues in subtilisin Novo are Pro210 and Gly211, with a normal *trans*-peptide bond between them. The ϕ–ψ angles for Gly211 in subtilisin Novo are 70° and 40°, not an unusual conformation for a glycine.

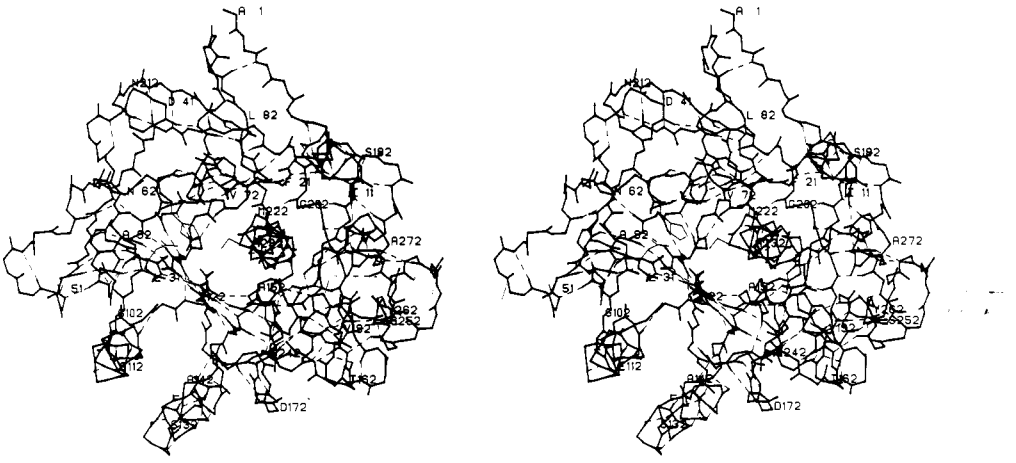

FIGURE 4: Subtilisin Carlsberg main-chain atoms and the hydrogen bonds between them (dashed lines). The active site is at the front of this view, and every 10th residue is numbered. Side chains are shown for the catalytic triad, Asp32, His64, and Ser221.

regions (Ramakrishnan & Ramachandran, 1965), with major clusters in the α-helix and β-sheet regions. The residues other than glycines that are outside the allowed regions (11 in EC–SC, 10 in CI-2–SN) are mainly found in irregular turns or segments of unclassified structure, often in regions of weak electron density. The final electron density maps calculated for the complexes are generally of excellent quality (see Figure 3). Some regions of poor density remain; all of the residues with weak density are on the surfaces of the enzymes or the inhibitors, and many have side chains protruding out into the solvent. For detailed information on the structure solution, refinement, error analysis, and quality of both complex structures, please refer to McPhalen (1986).

Discussion

The fully refined models of the EC–SC and CI-2–SN complexes have acceptable stereochemistry and correspondingly low R factors, indicating good agreement between the structural models and experimental observations. These refined high-resolution structures have been analyzed for details of hydrogen bonding, secondary structure, geometry of the enzyme active sites, and interactions between the enzymes and the inhibitors.

Subtilisin Structures. Subtilisin is a globular protein that is approximately heart-shaped. The active site is located in a relatively shallow depression on the surface of the molecule. The basic folding of the Carlsberg molecule is the same as that described for the unrefined Novo (Wright et al., 1969; Drenth et al., 1972) (Figure 4). It is divided into two domains by a central seven-stranded parallel β-sheet (Table II, Figure 5). The larger domain contains a bundle of seven α-helices packed against one face of the β-sheet, with five of the helices running approximately antiparallel to the strands of the sheet. The smaller domain contains the remaining two α-helices, also

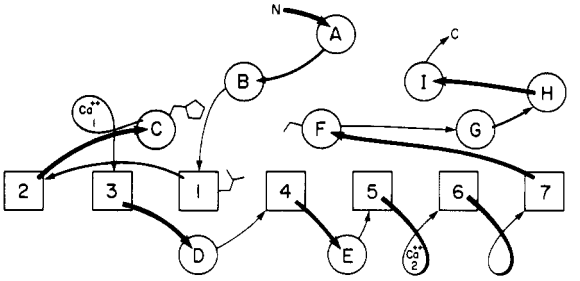

FIGURE 5: Connectivity and packing of the strands of the central parallel β-sheet (squares) and the α-helices (circles) in subtilisins Carlsberg and Novo. Labels correspond to those used in Table II. Helices C and F are essentially buried; the remainder are amphipathic and lie on the surface of the molecule. The approximate position of the active site is indicated by the side chains of the catalytic triad; the calcium binding loops are marked with Ca^{2+}.

Table II: Secondary Structural Elements of Subtilisin Carlsberg[a]

element	residues involved
parallel β-sheet	(1) Val26–Asp32, (2) Gly46–Phe50, (3) Ser89–Val95, (4) Asp120–Met124, (5) Val148–Ala153, (6) Ile175–Val180, (7) Val198–Gly202
antiparallel β-sheet	Val205–Tyr209, Thr213–Leu217
β-bridges[b]	Gly128, Tyr167, Ala179–Asp181, Asn185–Ala187, Thr255, Gly266–Ile268
α-helices	(A) Tyr6–Ile11, (B) Ala13–Ala18, (C) Gly63–Ala74, (D) Ser103–Asn117; (E) Ser132–Arg145, (F) Ser224–His238, (G) Ser242–Ser252, (H) Ser259–Gly264, (I) Asn269–Ala274
helical turn	Thr220–Ser224
type I turns[c]	Pro9–Lys12, Gln36–His39, His39–Leu42, Ala85–Val88, Asn97–Gly100, Tyr143Gly146, Tyr171–Val174, Asp181–Ser184, Ala187–Ser190, Gly193–Leu196, His238–Leu241, Ala272–Gln275
type II turns	Gly23–Val26, Val51–Glu54
type II′ turns	Ser159–Thr162, Tyr263–Gly266
type III turns	Pro5–Ile8, Val16–Gln19, Gln17–Gly20, Pro168–Tyr171, Gly219–Met222, Thr220–Ala223, Ala223–His226, Ser224–Val227, Ser259–Tyr262

[a] The elements of substilisin Novo are essentially identical. Due to bifurcation or irregularity in hydrogen-bond patterns, some residues participate in more than one secondary structural element. [b] β-structure involving only two hydrogen bonds (Kabsch & Sander, 1983). [c] The turn type is assigned on the basis of the ϕ_2–ψ_2 and ϕ_3–ψ_3 angles (Crawford et al., 1973).

packed antiparallel to the β-sheet. For those helices in subtilisin that pack one against the other, the average crossing angles are as follows: B/F, +35°; C/F, −35°; D/E, +14°; G/F, +116°; I/F, −31°; G/I, +129°. According to the protein taxonomy of Richardson (1981), subtilisin has a doubly wound parallel α/β structure. Both Carlsberg and Novo contain a highly unusual left-handed $\beta\alpha\beta$ crossover connection between strands 2 and 3 of the parallel β-sheet (Richardson, 1976). In addition, they contain a rare interrupted helix (Richardson, 1981); the C-terminal residue of helix C is Ala74 but, on the basis of hydrogen-bonding parameters and ϕ–ψ angles, residues Val84 and Ala85 can be considered a continuation of the helix. The segment of chain that interrupts the helix is the loop Leu75 to Gly83 that forms part of Ca^{2+}-binding site 1 (see below).

Subtilisin Carlsberg contains two *cis*-peptide bonds (McPhalen, 1986). The first is between Tyr167 and Pro168, corresponding to a *cis*-peptide found in the same position in Novo. The second is between Pro210 and Thr211 (Figure 3). Both of the *cis*-peptides in Carlsberg, as well as the corresponding bonds in Novo, are located between residues $i + 1$ and $i + 2$ of 5-membered reverse turns on the surface of the molecule. *cis*-Peptide bonds preceding proline residues have been observed for about 25% of all proline residues in protein crystal structures; they are found much more rarely before other residues due to energetic and conformational restrictions (Ramachandran & Mitra, 1976). Only four *cis*-peptide bonds preceding residues other than proline have been reported previously in reliable protein crystal structures (Creighton, 1983), although more are being seen as investigators come to believe that they are a rare but reasonable conformation.

Superpositions of Subtilisins Carlsberg and Novo. A global least-squares superposition (program of W. Bennett) of the 274 structurally equivalent C^{α} atoms in the two refined enzyme structures (omitting Thr55 of Novo) gives an rms deviation of 0.53 Å, a relatively small value that indicates substantial similarity between the two structures. Seventeen C^{α}'s differ in position by more than 1.0 Å, two of these by just over 2.0 Å (53 and 159): residues 16–20, 52–56, 131, 132, 159–162, and 211.

Residues 16–20 are part of helix B on the surface of the enzyme, far from the active site. The difference in the positions of helix B in the two enzymes is most likely due to changes in the amino acid sequences in this region between Carlsberg and Novo, resulting in differences in helix packing (Figure 6). Residues 52–56 are part of the loop containing Thr55 in Novo (deleted in Carlsberg). This is an exterior loop with relatively high *B* factors. Although the deletion of Thr55 is accompanied by sequence changes in six of the eight residues flanking it, the differences in the positions of the main-chain atoms are restricted to residues 52–56 (Figure 7). Residues 131 and 132 are part of an exterior loop that changes position in concert with a change in the position of the side chain of Tyr104. The different tyrosine side-chain positions may result from contacts with the P_4 side chain of the bound inhibitors (see on Enzyme–Inhibitor Interactions). Residues 159–162 are part of poorly ordered external loops in both enzymes. Residue 211 is the threonine preceded by a *cis*-peptide bond in Carlsberg; the equivalent residue is a glycine in Novo.

There are 82 amino acid differences between subtilisins Carlsberg and Novo (Figure 1a). The effects of these sequence changes on the structures of the enzymes were investigated by examining the two molecules overlapped with the global orientation matrix. About 40% of the differences are in residues that lie on the surface of the molecule and protrude into the solvent; they have a negligible effect on the course of the protein backbone. Another 15% are in surface crevices or the protein interior, but the changes are conservative and the side chains occupy essentially the equivalent volume in each enzyme. Several changes are less conservative in terms of the space occupied by the side chains, but hydrogen-bonding atoms overlap closely, e.g., the substitutions (Carlsberg to Novo) Thr33 → Ser33, Thr78 → Ser78, Ser183 → Asn183, and Ser242 → Thr242.

In several cases, a single amino acid change from Carlsberg to Novo is nonconservative but is compensated by another change close by in space; the net effect of the changes is that two side chains together occupy the equivalent volume in each enzyme. Examples of this are Tyr57 and Asn58 in Carlsberg occupying the equivalent space of Asn56 and Phe58 in Novo (Figure 7), Asn97 and Ser99 to Gly97 and Asp99, Tyr143 and Ala243 to Val143 and Asn243, Glu197 and Ser251 to Asp197 and Gln251, and Leu241 and Arg249 to Trp241 and Ser249. A somewhat more elaborate compensatory substitution is that of Gly63, Tyr209, and Leu217 in Carlsberg for Ser63, Leu209, and Tyr217 in Novo (Figure 8). Residue 217 forms part of

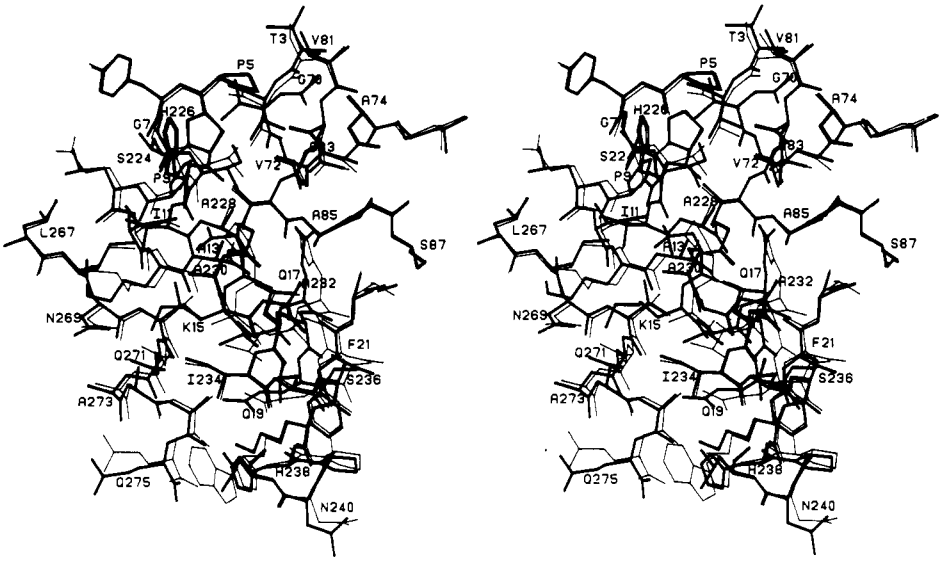

FIGURE 6: Major conformational difference in helix B of the two subtilisins (thick lines correspond to Carlsberg, thin lines to Novo). The enzymes were superposed by the global orientation matrix. Residues of helix C and part of the Ca^{2+} binding loop (Gly70–Ser87), helix F, and part of helix I superimpose closely. Residues Gly7–Phe21 occupy different positions. Amino acid substitutions at positions 8 (Val to Ile), 14 (Pro to Asp) and 16 (Leu to Val) from Novo to Carlsberg allow for a different placement of helix B against the protein core. The differences are as large as 1.5 Å in the region of residues 17–19.

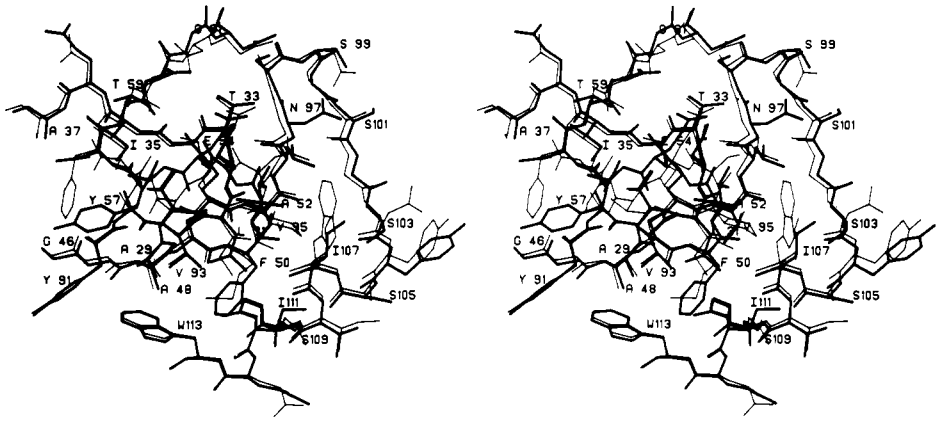

FIGURE 7: Conformational differences in the region of the deletion at Thr55 (thick lines are residues of Carlsberg; thin lines represent Novo). The strands of polypeptide in Carlsberg and Novo underlying the surface loop (Ala29–Ala37, Tyr91–Val95, Gly46–Phe50, and Ile107–Trp113) agree extremely well in position. Note also the compensatory effects of the amino acid differences Ala56, Tyr57, and Asn58 in Carlsberg and Asn56, Pro57, and Phe58 in Novo and how the residues' side chains occupy equivalent volumes. Main-chain atoms of residues 52–56 differ in position by over 1.0 Å; those of 57 and 58 differ by less than 0.4 Å.

the S_2' binding site (McPhalen et al., 1985b). Altogether, only localized structural changes are correlated with amino acid differences between Carlsberg and Novo, with the exception of the positional change in helix B. Similar observations on the "plasticity" of protein structures have been made for proteins altered by limited site-directed mutagenesis (Matthews, 1987).

Active-Site Regions of Subtilisins Carlsberg and Novo. The active sites of Novo and Carlsberg may be compared after functional groups in the region are overlapped (Figure 8, Table III). Even residues not directly involved in substrate binding and catalysis have almost identical conformations, and the rms deviation between the two sets of atoms (0.19 Å) is close to the estimated error in atomic positions for the Novo structure. After superposition of the active sites, the main-chain atoms of the loop from Leu96 to Ser105 still differ by 0.6–1.3 Å. Considering how closely the rest of the active-site regions overlap, this is a significant difference that points to an adjustment of this part of the active site to the conformation of different inhibitors. The loop lies closer to the inhibitor in Carlsberg than in Novo. In the crystal structure of the native Novo, this loop was described as very mobile and flexible

Table III: Hydrogen Bonds of the Catalytic Triad

bond[a]	N···O (Å)	H···O (Å)	N–Ĥ–O (deg)
His64 $N^{\epsilon 2}$			
Novo, Ser221 O^{γ}	2.73		154
Carlsberg, Ser221 O^{γ}	2.68		165
chymotrypsin,[b] Ser195 O^{γ}	2.51		175
His64 $N^{\delta 1}$			
Novo, Asp32 $O^{\delta 1}$	3.11	2.20	150
Novo, Asp32 $O^{\delta 2}$	2.79	1.83	159
Carlsberg, Asp32 $O^{\delta 1}$	3.14	2.27	145
Carlsberg, Asp32 $O^{\delta 2}$	2.68	1.71	162
chymotrypsin, Asp102 $O^{\delta 1}$	2.76	1.78	168
Asp32 $O^{\delta 1}$			
Novo, 0352 O	2.72		
Carlsberg, 0334 O	2.64		
chymotrypsin, Ser214 O^{γ}	2.61	1.61	175
Asp32 $O^{\delta 2}$			
Novo, Ser33 NH	2.71	1.79	153
Carlsberg, Thr33 NH	2.69	1.81	145
chymotrypsin, Ala56 NH	2.76	2.29	107
chymotrypsin, His57 NH	2.85	1.93	151

[a] His64 and Asp32 are equivalent to His57 and Asp102 in α-chymotrypsin. [b] The values for α-chymotrypsin are from its structure in complex with OMTKY3 (Read et al., 1984; Fujinaga et al., 1987).

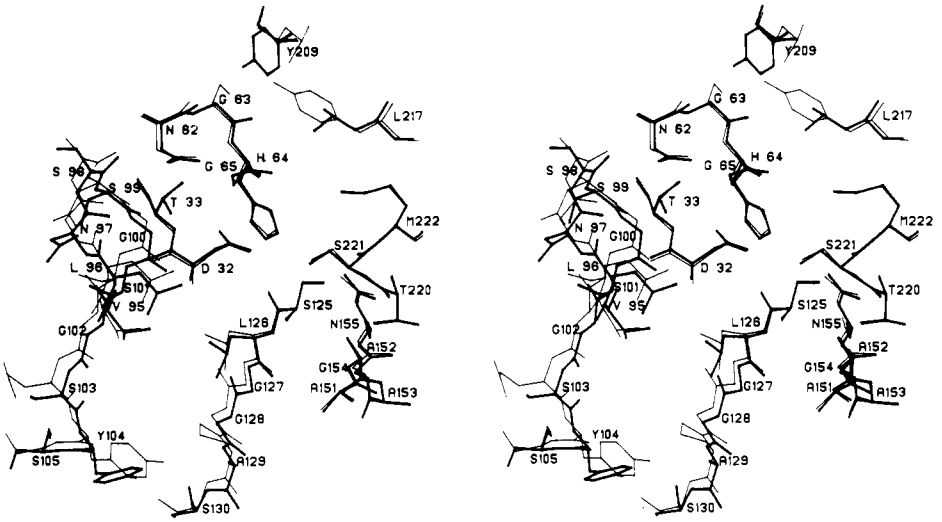

FIGURE 8: Superposition of active sites of subtilisins. All main-chain and C^{β} atoms (except glycines) of Leu31–Gly34, Asn62–Thr71, Met124–Gly128, Ala151–Asn155, and Leu217–His225 were overlapped. These segments include sequences flanking the catalytic triad and residues forming the walls of the S_1 specificity pocket and the oxyanion hole. The rms deviation in position of these 161 superposed atoms is 0.19 Å. Subtilisin Novo is drawn with thin lines, Carlsberg with thick lines. The residues of Carlsberg are labeled.

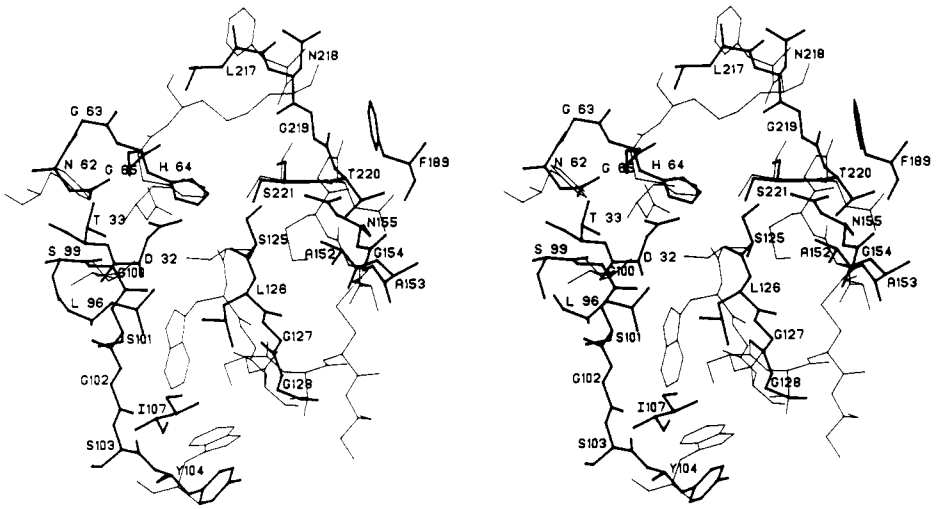

FIGURE 9: Superposition of subtilisin Carlsberg and α-chymotrypsin active sites. The active sites of the enzyme were superposed by a least-squares fit of the side chains of the catalytic histidine and serine residues, C^{α} of the histidine and C^{α} and N of the serine, plus the main-chain atoms of the P_1 and P_1' residues of the respective inhibitors. The rms deviation of the 19 superposed atoms is 0.46 Å. α-Chymotrypsin is drawn with thin lines, subtilisin Carlsberg with thick lines. The residues of Carlsberg are labeled.

(Drenth et al., 1972; Wright et al., 1969). The different positions of the loop in the two enzyme–inhibitor complexes may indicate some degree of induced fit in the formation of the complex. The position of the loop may also be affected by two sequence changes between the enzymes. The substitution of Thr33 (Carlsberg) for Ser33 (Novo) introduces an extra buried methyl group next to Leu96 as a possible source of steric crowding and conformational change (Figure 7). The deletion of Thr55 in Carlsberg is associated with small concerted positional changes of several segments of polypeptide chain, including the loop from Leu96 to Ser105 (Figure 7).

Overall, the few changes in the active-site region may have some effect on the specificity or strength of substrate–inhibitor binding, but the general similarity of the catalytic apparatus and binding sites indicates that there should be little difference in catalytic activity between Carlsberg and Novo. Thus, the observed differences in the ratios of esterase to proteinase activity (Barel & Glazer, 1968) must be due to more subtle conformational differences than those discussed here or to other factors such as differences in the electrostatic environment of the two active sites, Russell and Fersht (1987) have shown that site-specific mutations of charged residues (positions 99 and 156) that are within 15 Å of the catalytic triad of Novo to the corresponding neutral residues of Carlsberg can affect catalytic activity. The mutant enzymes showed increased pK_a values for the His64, Asp32 pair, altered k_{cat} values, and increased k_{cat}/K_M values for substrates.

Subtilisin Carlsberg and α-Chymotrypsin. Previous comparisons of the active-site regions of the native subtilisin Novo and α-chymotrypsin concluded that the geometry of the catalytic triad was virtually identical in the two enzymes (Wright, 1972; Kraut et al., 1972), although the overall folding of the molecules is different. A similar comparison has been made between subtilisin Carlsberg from the complex with eglin-c and a refined α-chymotrypsin structure from the complex with OMTKY3 [R factor = 0.168 for data 8.0–1.8 Å (Read et al., 1984; Fujinaga et al., 1987)] (Figure 9). The α-chymotrypsin from an enzyme–inhibitor complex was chosen to minimize possible differences in the conformation of catalytic residues as a result of substrate–inhibitor binding. The side chains of the two active-site histidines and serines, His64 and Ser221 in Carlsberg and His57 and Ser195 in α-chymotrypsin, superpose well. The active-site aspartates, Asp32 and Asp102, are quite different in position and orientation (Figure 9). The

angle between the planes of the two carboxyl groups in the overlapped proteins is 60°, and their hydrogen-bonding interactions are quite dissimilar (Table III). Other comparisons between the active sites pertain to enzyme–inhibitor interactions and are discussed below.

Solvent Structures. The final refined model of EC–SC includes 167 solvent molecules; 164 of these have been refined as water molecules and the remaining 3 as Ca^{2+} ions. CI-2–SN also contains 167 solvent molecules, including 2 refined as Ca^{2+} ions, but the solvent structure in the 2 complexes is not identical. Calculation of the amount of solvent to be expected in the asymmetric units of the crystals indicates that only about 15% of the solvent has been modeled. This low percentage is due in part to the highly conservative criteria used for choosing and retaining less ordered solvent in the models. A number of solvent molecules are conserved between the CI-2–SN and EC–SC complexes. If the complexes are superposed by the global orientation matrix of enzyme α-carbon atoms, 45 water molecules overlap within 1.0 Å. Some of the conserved waters are of special functional importance to the enzyme, the inhibitor, or the complex; these are described in the appropriate following sections. From the global superposition, 33 water molecules are conserved in crevices on the enzyme surfaces. Six of these lie in a crevice close to Leu75 and the loop forming the binding site of Ca^{2+} ion 1 (see below). An acetone molecule from the crystallization medium was observed in this region of the native subtilisin Novo (Drenth et al., 1972; Hol, 1971). Overlaps of individual secondary structural elements of the molecules indicate that the water structure on the surfaces of the complexes is more highly conserved than indicated by the global superposition. For example, superposition of helices D plus E of the two enzymes reveals 10 additional waters bound to the helices in equivalent positions.

The majority of the ordered water molecules lie on the surfaces of the complexes or in crevices in contact with other water molecules. Waters with less than 1% of their surface area accessible to bulk solvent (Lee & Richards, 1971) can be considered buried; 19 water molecules meet this criterion in Carlsberg and 16 in Novo. Several of the buried water molecules conserved between the two enzymes are located behind the active sites, interacting with polar side chains buried in the hydrophobic protein interior. The side chains of His67, Thr71, and His226 form hydrogen bonds with a channel of eight well-ordered waters that leads to the surface near the N-terminus of the enzyme. One buried water molecule is hydrogen-bonded to the side chains of Asn123 and the catalytic Asp32 (Table III).

Two monovalent cation-binding sites have been described previously for subtilisin Novo (Drenth et al., 1972). During the course of the structure refinement of the present subtilisin–inhibitor complexes, three solvent sites in the Carlsberg enzyme and two in Novo refined rapidly to occupancies of 1.0 and low *B* factors when their scattering contributions were included as water oxygen atoms (McPhalen, 1986). Subsequently, all of these ion sites were refined as Ca^{2+} ions [18e, scattering factor of Cromer and Mann (1968)] with no restraints on the Ca^{2+} to O ligand distances. The final occupancies and *B* factors of the three Carlsberg ions are 1.0 and 8 Å^2 for ion 1, 0.91 and 10 Å^2 for ion 2, and 0.88 and 13 Å^2 for ion 3. Those for Novo are 0.97 and 10 Å^2 for ion 1 and 0.72 and 28 Å^2 for ion 2. Subtilisin Novo is known to be stabilized by Ca^{2+} (Matsubara et al., 1958). Many other extracellular proteinases are stabilized by Ca^{2+} ions as a protection against autolytic degradation (Kretsinger, 1976; Bode & Schwager, 1975; Read & James, 1988). Any calcium ions present in these structures of Carlsberg and Novo presumably were copurified with the proteins, since no calcium was added to the crystal growth solutions.

Table IV: Ion Coordination Geometry in Subtilisins

site	ion ligands	ligand position	mean Ca^{2+}···O distance (Å)	mean O–C–O angle[a] (deg)
1	Gln2 O^{ϵ_1}, Asp41 O^{δ_1},O^{δ_2}, Asn77 O^{δ_1},Val81 O	equatorial	2.4	73
	Leu75 O,Thr79 O	axial	2.4	91
2	Carlsberg			
	Tyr171 O, Val174 O, 2 H_2O	equatorial	2.6	86
	Ala169 O	axial	2.6	106
	Novo			
	Tyr171 O, Val174 O, Asp197 O^{δ_2},Glu195 O	equatorial	2.9	67
	Gly169 O, 1 H_2O	axial	2.9	91
3	Ala37 O, His 39 O, Leu42 O, 2 H_2O		2.8	

[a] For equatorial ligands, measured with respect to neighboring equatorial ligands. For axial ligands, measured with respect to all equatorial ligands.

The coordination geometry of the ions is given in Table IV, and stereoviews of some of the ion-binding sites are shown in Figure 10. In both Carlsberg and Novo the high occupancies, low *B* factors, and mean ion to oxygen ligand distance of 2.37 and 2.40 Å, respectively, are strongly suggestive that ion site 1 is occupied by a Ca^{2+} ion (Einspahr & Bugg, 1981, 1984). A least-squares superposition of equivalent atoms in the several enzyme residues that are Ca^{2+} ligands [all atoms of Gln2, Asp41, Leu75, Asn77, Val81, and the main-chain and C^{β} atoms of Thr79 (Ile79 in Novo)] gives an rms deviation of 0.14 Å. The two Ca^{2+} ion positions agree to within 0.12 Å, significantly smaller than the rms coordinate error of subtilisin Novo. This Ca^{2+} binding loop is remarkably similar in both enzymes but has been deleted in the structurally similar proteinase K (Jany et al., 1986).

The other two ion-binding sites in Carlsberg are less well conserved in Novo. Site 2 in Novo has weaker electron density and a mean ligand distance of 2.91 Å, more typical of K^+···O (Brown & Shannon, 1973) or H_2O···O distances (Table IV). The geometry of this site in Carlsberg is that of a tetragonal pyramid, whereas in Novo it is a pentagonal bipyramid with one empty equatorial site. Site 3 in Carlsberg has a relatively high occupancy and low *B* factor, but the mean ligand distance of 2.85 Å is again more typical of a K^+ ion. In Novo the equivalent position is occupied by a well-ordered water molecule (occupancy = 1, *B* factor = 14 Å^2) with three protein ligands.

Inhibitor Structures. The eglin-c and CI-2 molecules are oblate ellipsoids of revolution and can be described as wedge-shaped disks (Figure 11), similar in size and shape to the avian ovomucoid inhibitors (Fujinaga et al., 1982; McPhalen et al., 1985b; Read & James, 1986). The reactive-site loop forms the narrow end of the wedge. The reactive-site bond is Leu59I–Asp60I in eglin-c and Met59I–Glu60I in CI-2 (residue numbering as in Figure 1b). The polypeptide chain fold of eglin-c is similar to that of CI-2, but both are different from the tertiary organization of the ovomucoid inhibitors and other known inhibitor families (Marquart et al., 1983; Hirono et al., 1984; Read & James, 1986). The assignment of secondary structural elements in the inhibitors has changed somewhat from our initial structure reports (McPhalen et al., 1985a,b); further refinement has al-

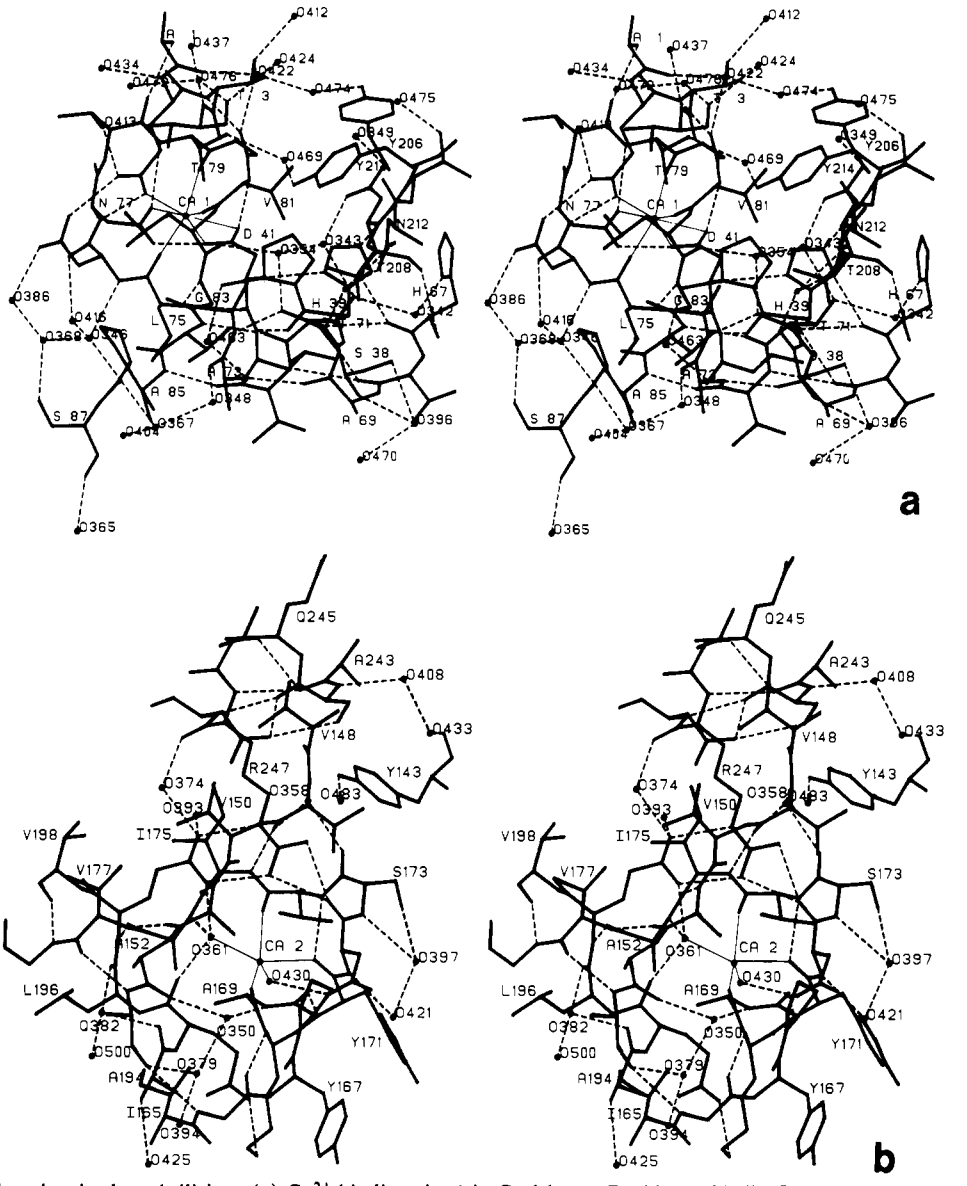

FIGURE 10: Ion-binding sites in the subtilisins. (a) Ca^{2+} binding site 1 in Carlsberg. Residues of helix C are seen at the bottom of the figure. A loop from Ala74 to Gly83 extends out from the helix to form the binding site. Four residues from this loop plus the side chains of Asp41 and Gln2 donate the oxygen ligands; there are no water molecules in the ligand field. This site is heptacoordinate (continuous thin lines) and is a slightly distorted pentagonal bipyramid (Table IV). Hydrogen-bonding interactions (dashed lines) provide additional stabilization of the loop conformation. (b) Ca^{2+} binding site 2 in Carlsberg. The average ion to oxygen distance would indicate that it is intermediate between a Ca^{2+} site and a K^+ site. It is approximately a tetragonal pyramid with a rectangular base (Table IV).

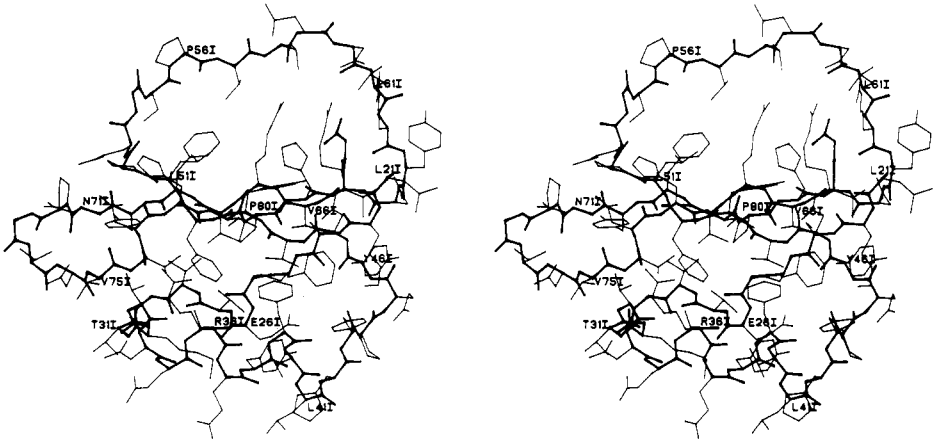

FIGURE 11: All non-hydrogen atoms of eglin-c, including only residues seen in the electron density map of the EC–SC complex, Leu21I–Gly83I. The main chain is drawn with heavy lines, side chains with light lines. Every fifth amino acid residue is labeled.

tered some hydrogen-bond assignments. The secondary structural elements of CI-2 and eglin-c are very similar, except in the loop containing the insertion of Gly72A in eglin-c (Table V).

Table V: Secondary Structural Elements of Eglin-c and CI-2

element	eglin-c	CI-2
parallel β-sheet	Asn47I–Glu53I, Asn64I–Tyr70I	Gln47I–Val53I, Asp64I–Val70I
antiparallel β-bridges	Lys22I–Phe24I, Lys30I–Val32I, Asn74I–Val76I, Pro80I–Val82I	Thr22I–Trp24I, Lys30I–Val32I, Phe69I–Asp71I, Asp74I–Ala77I, Pro80I–Val82I
α-helix	Thr31I–Tyr43I	Ser31I–Lys43I
type I turns	Pro25I–Val28I, Tyr43I–Tyr46I, Arg62I–Arg65I	Lys43I–Ala46I, Arg62I–Arg65I, Asp71I–Asp74I
type II turns	Val27I–Lys30I, Pro52I–Ser55I	Leu27I–Lys30I
type III turn	Phe24I–Val27I	Trp24I–Leu27I
unclassified hydrogen bonds between main-chain atoms	Gly29I–Val76I, Arg62I–Gly83I, Asn71I–Val75I, Asn74I–Asn71I	Gly29I–Ile76I, Arg62I–Gly83I, Gly83I–Arg65I

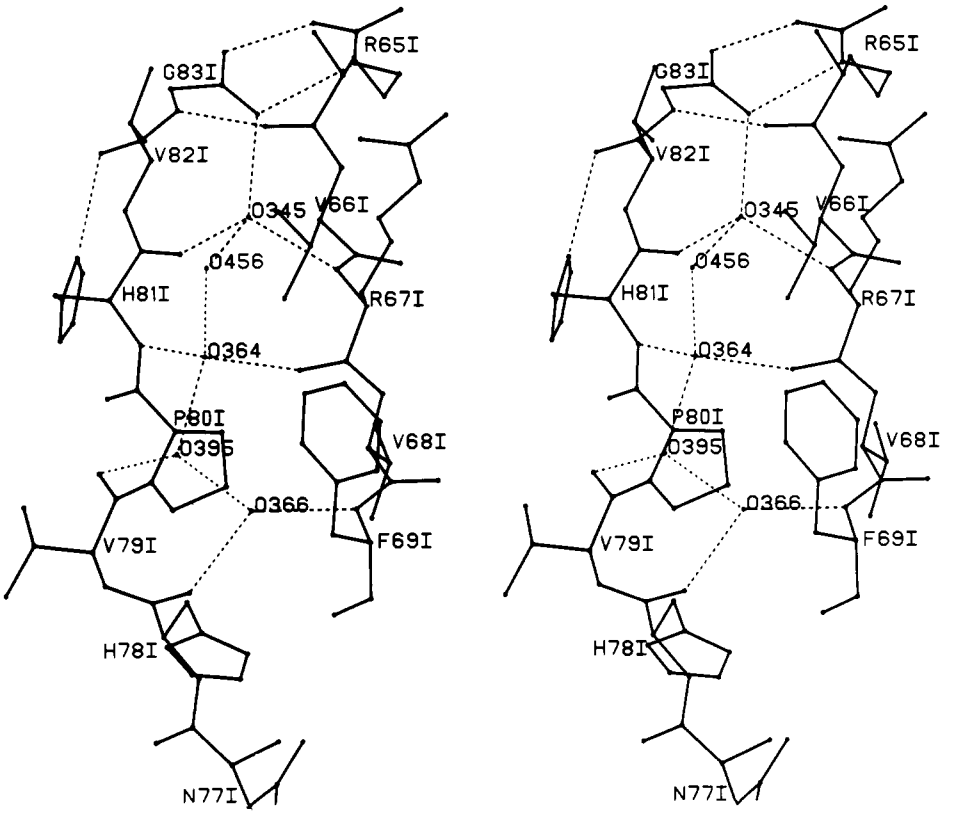

FIGURE 12: Bridging water molecules in the β-sheet of eglin-c. Strand 2 of the parallel β-sheet and the C-terminal strand of eglin-c are shown, with the water molecules providing hydrogen-bonding bridges (dashed lines) between them.

The body of each inhibitor is composed of an α-helix of 13 residues, packed against an irregular four-stranded mixed parallel and antiparallel β-sheet. The sheet in eglin-c is made up of Lys22I–Phe24I at one edge, with two antiparallel hydrogen bonds to the C-terminal strand, and strands 1 and 2 of the parallel β-sheet. The C-terminal strand is connected to strand 2 only by a set of well-ordered water molecules participating in linking hydrogen-bonding bridges (Figure 12). These strands in CI-2 are linked by conserved water molecules in the same way, as well as by two additional hydrogen bonds between main-chain atoms (McPhalen, 1986). The conservation of these waters in both complexed inhibitors, as well as in the more flexible free CI-2 (McPhalen & James, 1987), indicates that they play an important role in the structural integrity of this family of inhibitors. The arrangement of the four β-strands has the characteristic twist of β-sheets, and the α-helix lies in the curvature created by the twist, with the helix axis antiparallel to the direction of the two parallel strands. The hydrophobic core of the molecule is the interface of the sheet and the helix, formed by residues Phe24I, Val27I, Val32I, Ala35I, Phe39I, Val66I, Val68I, Val76I, Pro80I, and Val82I (eglin-c sequence). All these residues have less than 10% of their surface area accessible to bulk solvent (Lee & Richards, 1971; Shrake & Rupley, 1973). Other partially buried residues (<25% accessible) contribute to the hydrophobic core: Tyr38I, Tyr43I, Tyr46I, Val48I, Phe50I, Leu51I, and Tyr70I. Being small proteins, both inhibitors have a similar number of hydrophobic residues partially or fully exposed on their surface.

The broad loop containing the reactive-site bond lies on the opposite side of the β-sheet from the α-helix. The second strand of parallel β-sheet contains Arg65I and Arg67I, residues with extended side chains that provide part of the hydrogen-bonding network that supports the extended conformation of the loop (Figure 13). The other residues involved in the network for eglin-c are Thr58I, Asp60I, Arg62I, Gly83I, and one tightly bound water molecule (O372). The hydrogen-bonding capacity of all residues is fully used; 12 of the 15 possible hydrogen bonds of these residues form the network, while the remaining 3 are to the enzyme or other tightly bound water molecules. The three positive charges of the arginine side chains are balanced by Asp60I and the carboxylate of the C-terminal Gly83I; they are also partially accessible to bulk solvent.

Pyramidalization of the carbonyl carbon atom of the reactive-site bond indicates how far an enzyme-bound inhibitor

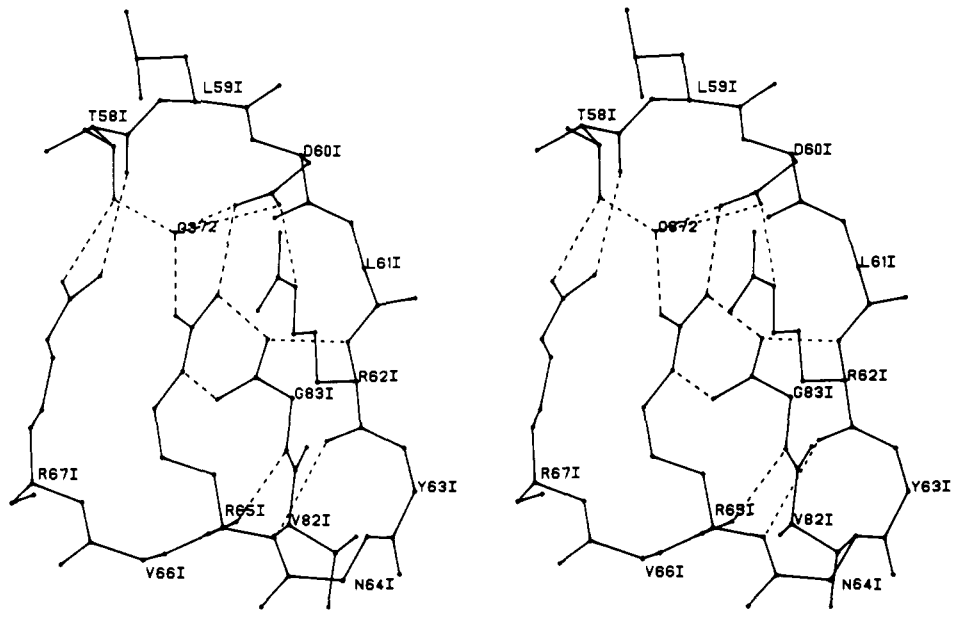

FIGURE 13: Hydrogen bonding of the reactive-site loop in eglin-c. Dashed lines indicate hydrogen bonds. Side chains of residues not involved in hydrogen bonding have been omitted for clarity.

has moved along the reaction pathway toward a tetrahedral intermediate. A review of many serine proteinase–inhibitor complexes (Read & James, 1986) concludes that although small distortions of this bond are observed, the size of the distortion is comparable to the coordinate errors of the structures. In eglin-c the torsion angle ω (C^{α}_i–C_i–N_{i+1}–C^{α}_{i+1}) is 170° (175° in CI-2), close to the expected value of 180°. The θ^4 angle (out-of-plane bending of the carbonyl oxygen with respect to C^{α}, C, and N) of Leu59I is −3.7° (−0.03 Å deviation from the plane), indicating that no detectable pyramidalization has taken place.

In spite of restraining the atoms of all of the peptide linkages to planarity (see Table I), larger deviations in θ^4 are seen in other residues of EC–SC [rms deviation 2.5°, corresponding to a displacement of the carbonyl carbon atom of 0.02 Å; the maximum θ^4 is at Asn184 (−8.6°, −0.07 Å)]. The corresponding values for CI-2–SN are rms deviation 1.9° (0.02 Å), maximum at Glu60I of 6.0° (0.05 Å). The values for the scissile peptide in CI-2 are 0.6° (<0.01 Å).

Superpositions of Eglin-c and CI-2. The structures of eglin-c and CI-2 were overlapped by a least-squares superposition of structurally equivalent α-carbon atoms, Lys22I–Pro72I and Thr73I–Gly83I. The rms difference in position for these 62 C^{α}'s is 1.68 Å. Iteration of the least-squares process and gradual removal of the most widely separated pairs of atoms from the calculation leads to an rms deviation of 0.73 Å for 39 of the 62 α-carbon pairs. Residues included in the calculation of this final rms deviation are Lys22I–Lys30I (eglin-c sequence), Gln45I–Glu53I, Leu59I–Asn71I, and Val76I–Gly83I. The regions of best agreement between the two inhibitors are thus the four strands of β-sheet and the second half of the reactive-site loop.

The α-helices overlap poorly. The residues of the α-helix show deviations of 1.9–5.2 Å between α-carbon atoms. The helix position differs by an angle of 18° and a translation of 0.8 Å between eglin-c and CI-2. In eglin-c, this helix has a number of packing contacts with symmetry-related molecules that do not occur in CI-2. These contacts may be partly responsible for the shifted helix position. However, most of the change in helix position appears to be due to differences in the packing of the hydrophobic core between eglin-c and CI-2. Almost 75% of the residues forming the core are different in the inhibitor sequences (Figure 1b). The core of CI-2 contains 1 aromatic residue and 15 aliphatics; eglin-c contains 7 aromatics and 10 aliphatics.

A model of a "mutated" CI-2 molecule was built by replacing the amino acid side chains of core residues by the corresponding side chains from eglin-c. The CI-2 model retained positions of main-chain atoms and the applicable side-chain torsion angles of the real CI-2 core residues (program of R. Read). The mutated CI-2 molecule was examined on the MMS-X graphics system. Bad packing contacts in such a model could indicate whether conformational changes of core residues are necessary to accommodate the eglin-c sequence. Three bad contacts were found, all of which could be relieved by a shift of the eglin-c helix in the direction of the observed difference. The bad contacts resulted from the substitution of Tyr46I (eglin-c) for Ala46I (CI-2), Phe50I for Val50I, and Phe39I, for Ile39I. The different position of the helix of eglin-c also allows a more favorable stacking of the side-chain aromatic ring of Tyr38I with those of Phe39I, Phe24I, and His42I.

A second segment of eglin-c that deviates in position from CI-2 in the global superposition of α-carbon atoms is Gly54I–Thr58I, the first part of the reactive-site loop (2.1–4.1-Å difference in α-carbon atoms). The enzyme residues interacting with this portion of the reactive-site loop, Gly100–Ser105, are also in different positions between the active sites of subtilisins Carlsberg and Novo (Figure 8). The reactive-site loops of eglin-c and CI-2 were overlapped by a least-squares superposition of the main-chain plus C^{β} atoms of the P_3 to P_1' residues and the main-chain atoms of P_4 (proline in eglin-c). The rms deviation of these 24 atoms was 0.19 Å (Figure 14). The very close overlap of the reactive-site loops, including the P_4 to P_1' residues, is in contrast to the differences seen in this region when the whole inhibitors are overlapped. The position of the body of the inhibitor relative to the reactive-site loop thus is different between the two complexes. A similar difference was seen in the structures of OMTKY3 in complex with SGPB and α-chymotrypsin (Read et al., 1984; Fujinaga et al., 1987) and attributed to a combination of changes in interactions with the different enzyme active sites and changes in crystal packing contacts. Both factors may apply to eglin-c and CI-2, and differences in

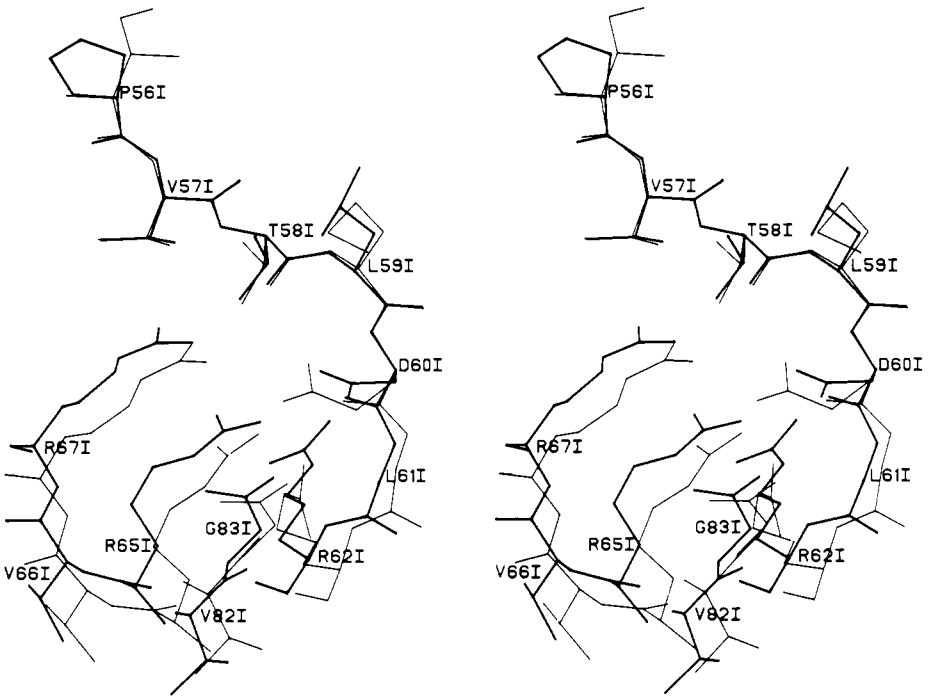

FIGURE 14: Superposition of eglin-c and CI-2 reactive-site loops. The reactive-site loops of the two inhibitors were overlapped as described in the text. CI-2 is drawn with thin lines, eglin-c with thick lines. The residues of eglin-c are labeled.

sequence of the reactive-site loops may also contribute to the observed conformational differences.

There is one major change between the two reactive-site loops, the substitution of Asp60I in eglin-c for Glu60I in CI-2. The P_1' side chain is extensively involved in hydrogen bonding that presumably stabilizes the conformation of the loop and the reactive-site bond (Figure 13). P_1' is Asp or Glu also in 99 of 100 sequences of avian ovomucoid inhibitors (Laskowski et al., 1987). The $O^{\delta 2}$ of Asp60I and the $O^{\epsilon 2}$ of Glu60I are 0.6 Å apart in the superposition (Figure 14); this is relatively close, since the Asp60I side chain is one carbon shorter than that of Glu60I. One factor compensating for the shorter P_1' side chain is a difference in χ_1 of 30° and in orientation of the two carboxyl groups. A second factor is a shift of the side chains of Arg62I, Arg65I, Arg67I, and Gly83I toward Asp60I relative to their positions in CI-2 (Figure 14). A third factor is the presence in eglin-c of an additional water molecule adjacent to Asp60I $O^{\delta 2}$ (Figure 13). This water molecule (O372) provides hydrogen-bonding bridges between Asp60I, Thr58I, and Arg65I. The total number of hydrogen bonds in the network and the atoms involved in the bonds are thus highly conserved between the two inhibitors. This is an indication of the importance of the network to the stability of the reactive-site loop.

Enzyme–Inhibitor Interactions. Contacts between each inhibitor and its cognate enzyme in the complex are limited almost entirely to residues in the reactive-site loop (McPhalen et al., 1985a). This phenomenon is common to most serine proteinase inhibitors of known structure (Read & James, 1986). Twelve residues from eglin-c make a total of 134 contacts of less than 4.0 Å with 25 residues from subtilisin Carlsberg (Table VI). Eleven residues of CI-2 make 134 contacts with 22 residues from Novo in the fully refined CI-2–SN structure. The regular interactions of inhibitors or substrate analogues bound to serine proteinases are seen in the EC–SC and CI-2–SN complexes (Figure 15). The side chain of the P_4 inhibitor residue (Pro56I in eglin-c, Ile in CI-2) fits into a hydrophobic pocket on the enzyme surface, bound by main-chain atoms of Gly127, Gly128, Ser101, and Gly102

Table VI: Contacts of Less than 4.0 Å between Subtilisin Carlsberg and Eglin-c

residue	site	no. of contacts[a]	comments
Tyr49I		3	hydrophobic, with Ser99 C^{β}
Leu51I		1	close contact with Ser101 O^{γ}
Gly54I	P_6	17 (1)	close contacts with Tyr104, Gly128, Ala129, Ser130 (exterior loop)
Ser55I	P_5	3	close contacts with Gly127, Gly128
Pro56I	P_4	15 (1)	β-bridge with Gly102; hydrophobic, with S_4 binding pocket (Gly127, Gly128, Ser101, Gly102, Tyr104 side chain, Ile107 side chain)
Val57I	P_3	13 (2)	β-bridge with Gly127, close contacts with Gly100, Ser101, Leu126
Thr58I	P_2	17 (1)	β-bridge with Gly100; hydrophobic, with S_2 binding pocket (Thr33, His64 side chain, Leu96, Gly100)
Leu59I	P_1	31 (2)	carbonyl O in oxyanion hole (Ser221, Asn155); 3.6 Å from NH to Ser125 CO; hydrophobic, with S_1 binding cleft (Leu126, Gly127, Ala152, Gly154, Asn155); close contacts His64, Asn155, Gly219, Thr220, Ser221
Asp60I	P_1'	12	close contacts with His64 Asn155, Asn218, Ser221
Leu61I	P_2'	8 (1)	β-bridge with Asn218; hydrophobic, with Phe189; close contacts with Asn155, Asn218
Arg62I	P_3'	10 (3)	side-chain hydrogen bonds to Asn62; close contacts with Asn64, His64, Tyr209
Arg65I		4	close contacts of side chain with Ser99, Gly100

[a] Number of hydrogen bonds included in number of contacts given in parentheses.

and the side chains of Tyr104 and Ile107. The enzyme residues forming the pocket are completely conserved between Carlsberg and Novo, but the pocket differs in size and shape between the two. In Carlsberg, Tyr104 takes up an alternate conformation, rotated 30° in χ_1 and 50° in χ_2 out toward the solvent (Figure 8). The pocket in Carlsberg thus accommodates the shorter, broader side chain of Pro56I more easily. Tyr104 is one residue that exhibits conformational flexibility

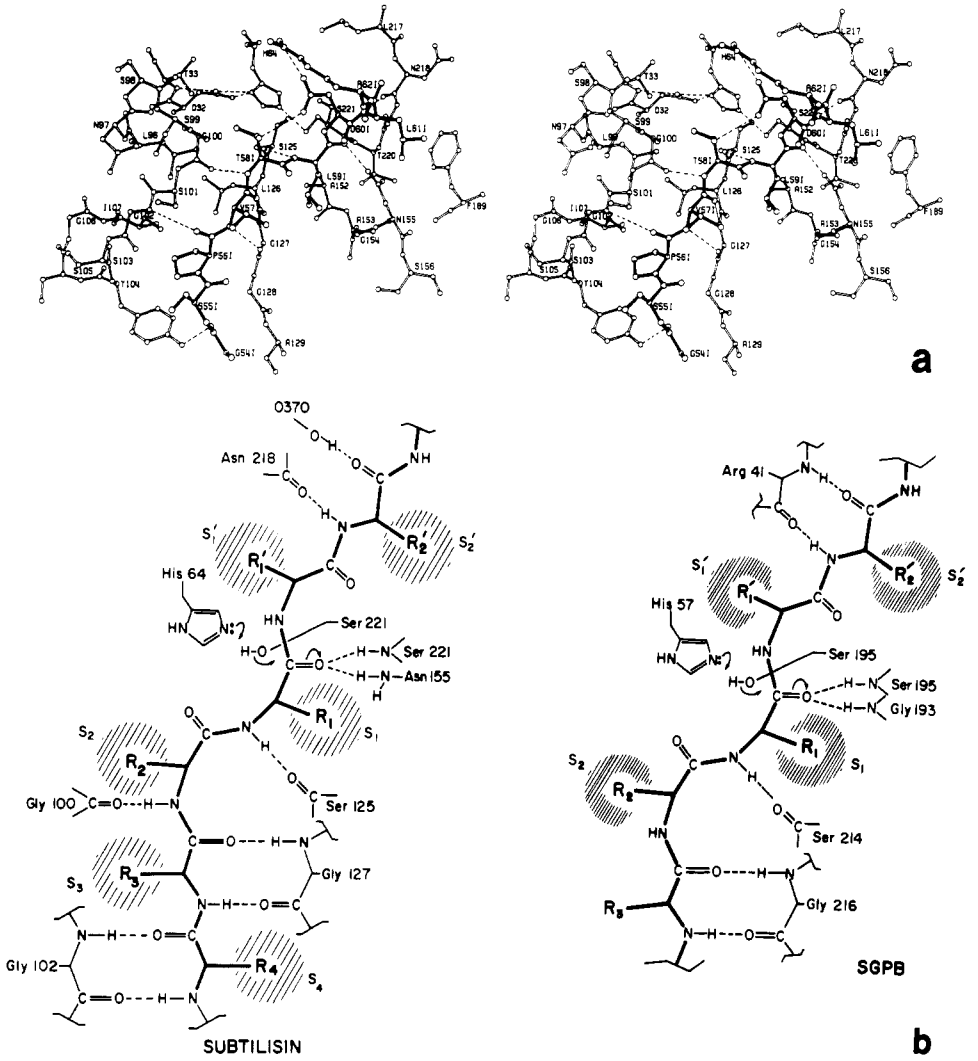

FIGURE 15: Interactions between subtilisin Carlsberg and eglin-c. (a) ORTEP drawing (Johnson, 1965) of the active-site region of Carlsberg (open bonds) and the reactive-site loop of eglin-c (filled bonds). Hydrogen bonds are indicated with dashed lines. The hydrogen bond shown from the NH of Leu59I (P_1) to the carbonyl oxygen of Ser125 is long (3.64 Å) and probably weak. This bond is characteristic of an incipient tetrahedral intermediate (Robertus et al., 1972; James et al., 1980). Part of the stabilization energy for the formation of a tetrahedral intermediate would be obtained by shortening, and thus strengthening, this hydrogen bond. (b) Diagrammatic representation of serine proteinase–inhibitor active-site interactions. Shown are specificity pockets and hydrogen-bonding interactions of the subtilisins (left) and of chymotrypsin-like enzymes, represented by SGPB (right).

among the many mutants of subtilisins made and studied crystallographically (David Estell and Rick Bott, personal communication).

The side chain of the P_3 residue is involved in hydrophobic interactions with the body of the inhibitor and has no specific binding site on the enzyme. The binding site for the side chain of the P_2 residue (Thr58I) is another hydrophobic pocket composed of the imidazole ring of His64, $C^{\gamma 2}$ of Thr33 and the side chain of Leu96 (Figure 15). The $O^{\gamma 1}$ of Thr58I is also involved in the hydrogen-bond network supporting the reactive-site loop in the bound inhibitor. The P_1 residue (Leu59I in eglin-c, Met in CI-2) lies in a relatively open cleft, bound by hydrophobic regions of Leu126, Gly127, Ala152, Gly154, and Asn155. There are additional hydrophobic interactions between Phe189 and the P_2' side chain (Leu61I in eglin-c, Tyr in CI-2).

The main-chain atoms of the P_1–P_4 residues in these inhibitors form the central strand of a three-stranded antiparallel β-sheet (Table VII). The flanking strands come from the enzyme, residues Ser125–Gly127 and Gly100–Gly102 (Figure 15). The third strand of β-sheet from the enzyme (100–102) is unique to the subtilisin family and was first described by Hirono et al. (1984) for the SSI–subtilisin Novo complex. A complete β-bridge for P_4 (P_4 NH to Gly102 CO and Gly102 NH to P_4 CO) is present only in CI-2–SN. The P_4 NH to

Table VII: Hydrogen Bonds (Å) in the Active Sites of Serine Proteinase–Inhibitor Complexes

site	Carlsberg–eglin-c	Novo–CI-2	SGPB–OMTKY3[a]	trypsin–PTI[b]
P_6 I CO–Tyr104 OH	2.5			
P_4 I NH–Gly102 CO		2.9		
P_4 I CO–Gly102 NH	3.4	3.0		
P_3 I NH–Gly127 CO	2.9	3.0	2.9	3.3
P_3 I CO–Gly127 NH	2.9	3.0	3.0	
P_2 I NH–Gly100 CO	2.9	3.1		
P_1 I NH–Ser125 CO	3.6	3.4	3.6	3.4
P_1 I CO–Ser221 NH	3.2	3.0	3.1	3.1
P_1 I C–Ser221 O^{γ}	2.8[c]	2.7	2.7	2.6
P_1 I CO–Asn155 $N^{\delta 2}$	2.8	2.7	2.6	2.9
P_2' I NH–Asn218 CO	2.8	2.8	2.8	
P_2' I CO–0370 O	2.6	2.7	3.2	3.0

[a] The equivalent residues from the subtilisins to SGPB and trypsin are Ser125 → Ser214, Gly127 → Gly216, Asn155 → Gly193 NH, Asn218 → Phe41, and Ser221 → Ser195. Data from Read et al. (1983). [b] Data calculated from coordinates obtained from the Brookhaven Protein Data Bank (Bernstein et al., 1977). [c] This is the contact distance from O^{γ} of the active-site serine to the carbonyl carbon atom of the scissile peptide.

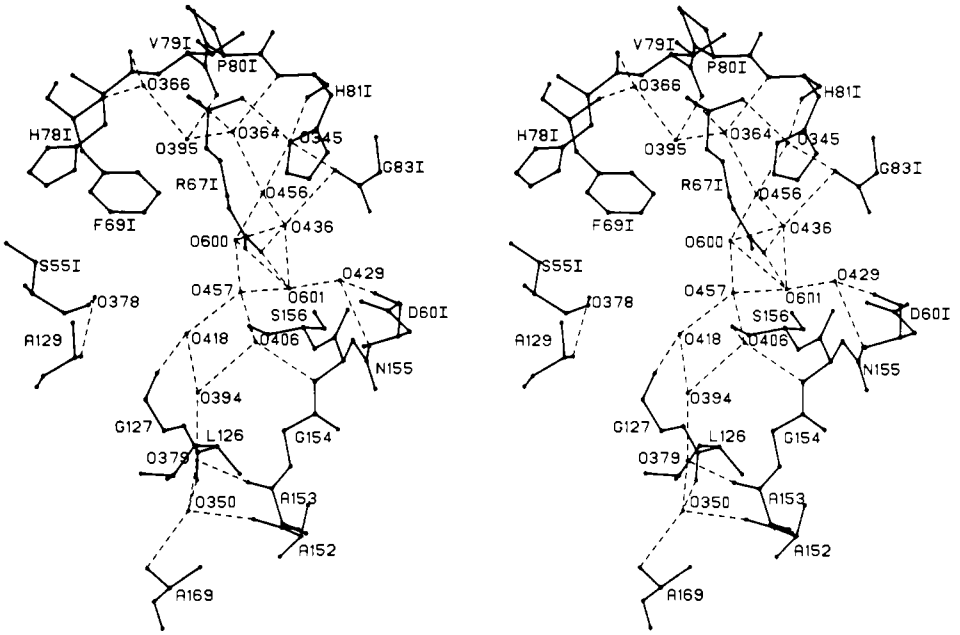

FIGURE 16: Water molecules at the interface of the EC–SC complex. Dashed lines indicate hydrogen bonds. Water molecules with relatively weak electron density have sequence numbers beginning with 600. Ordered waters in the S_1 binding subsite are shown.

Gly102 CO bond is missing in EC–SC (P_4 is a proline), and Gly102 NH to P_4 CO is missing in the SSI–subtilisin Novo complex (Hirono et al., 1984). The backbone of the loop from Gly100 to Tyr104 and the side chain of Tyr104 form part of the deep hydrophobic P_4 binding pocket in subtilisin; these residues may change position to accomodate different P_4 side chains (Pro in eglin-c, Ile in CI-2, Met in SSI) and thus change the hydrogen-bonding scheme in the different complexes.

Serine proteinase–inhibitor complexes have another important regulator interaction, seen also in the EC–SC and CI-2–SN complexes. The carbonyl oxygen of P_1 lies in the oxyanion hole, with hydrogen bonds from the main-chain NH of the catalytic Ser221 and the $N^{\delta 2}$ of Asn155 (Table VII). Development of an enhanced positive charge on the carbonyl carbon of P_1 through polarization of the CO bond is believed to be important in "inducing" the nucleophilic attack of Ser221 in catalysis (James et al., 1980). Thus, a close approach ($\simeq$2.7 Å) of Ser221 O^{γ} to the P_1 carbonyl carbon is observed in both complex structures (Table VII). One other feature in the subtilisins that may have an influence on inhibitor or substrate interactions is the helix dipole of helix F. The positively charged end of the dipole points at Ser221 and may enhance the catalytic action of subtilisin by promoting proton transfer (Hol, 1985). The helix dipole also provides an additional stabilizing interaction for the P_1 carbonyl oxygen in the oxyanion hole.

In chymotrypsin-like enzymes the hydrogen bond from Asn155 $N^{\delta 2}$ to the carbonyl oxygen of P_1 in the oxyanion hole is replaced by one from Gly193 NH (Table VII). The side chains of asparagine residues are normally less rigid in conformation than the nitrogen of a peptide bond. Asn155 is carefully positioned, however, by additional hydrogen bonds between its $O^{\delta 1}$ and $N^{\delta 2}$ and the NH (2.9 Å) and $O^{\gamma 1}$ (2.8 Å) of Thr220. The *B* factors of $O^{\delta 1}$ (10 Å^2) and $N^{\delta 2}$ (9 Å^2) reflect the relatively rigid side-chain orientation; the average values for the 18 other asparagines in Carlsberg are 23 Å^2 for $O^{\delta 1}$ and 19 Å^2 for $N^{\delta 2}$.

A number of well-ordered solvent molecules provide a network of hydrogen-bonding bridges between the inhibitors and the enzymes of these complexes (Figure 16). Similar ordered solvent structures have been described for other pairs of serine proteinase–inhibitor complexes (James et al., 1980; Fujinaga et al., 1987; Bode & Huber, 1978; Bode et al., 1987). In the EC–SC complex there are two deeply buried waters (O350 and O379) that are probably structural to the subtilisin molecules and would not be displaced even by a larger P_1 residue. They are analogous to three water molecules previously identified in the S_1 pocket of α-chymotrypsin (Fujinaga et al., 1987). Seven other waters (O394, O418, O406, O457, O600, O601, and O429) may or may not be displaced depending upon the nature of the P_1 residue. Waters O378 and O429 make direct hydrogen-bonded bridges between Carlsberg and eglin-c. There are fewer ordered waters in the vicinity of the S_1 site in the CI-2–SN complex. Two deeply buried in the S_1 pocket are analogous to the two in subtilisin Carlsberg, and one bridges the amide of Asn155 and the carbonyl of Glu60I analogous to O429. This water also forms a hydrogen bond to the carboxylate of Gly83I in the CI-2 inhibitor. Further direct hydrogen-bonded bridges are made by conserved water molecules between the inhibitor P_2, P_1, and P_2' residues and the S_3, S_2, and S_2' sites on the enzymes.

The role of ordered solvent molecules at the interface of enzyme and inhibitor is difficult to assess in energetic terms. These waters are not displaced upon the confrontation of the inhibitor and enzyme, a process that would provide a favorable entropic term. Many of these water molecules, however, especially those bridging between groups on the enzyme and the inhibitor, are contributing to the overall complementarity of the two interacting surfaces. It is clear that their contribution will have to be considered when computations of interaction energies in docking algorithms are designed and developed.

Another Carlsberg–Eglin-c Complex. Recently Bode et al. (1986, 1987) have presented an independent determination of the structure of the EC–SC complex, using the inhibitor isolated from leeches rather than the genetically engineered product. Although the crystals of Bode et al. were grown from 10% poly(ethylene glycol) solution instead of 0.7 M phosphate, the space group of both complexes is P_1 and the unit cell dimensions are almost identical. Bode et al. present some details of the enzyme and inhibitor structures that may be compared with ours. Comparisons of the overall chain fold and secondary structural elements suggest that the enzymes

in the two structures are close to identical. The assignments of secondary structural elements (Table II) differ in minor points, e.g., the exact residues at which an element begins or ends.

One important difference between the two enzyme structures is the nature of the peptide bond preceding Thr211. In our model, this is an unusual *cis*-peptide lying in the middle of a five-membered turn (Figure 3). The *B* factors of the main-chain atoms in this turn are all less than 13 Å^2, and the main-chain torsion angles of Thr211 are $\phi = -103°$ and $\psi = 120°$. Bode et al. presumably have interpreted this bond as a normal *trans*-peptide. In their Ramanchandran plot Thr211 is the only β-branched amino acid with ϕ and ψ near the left-handed α-helical regions. A sharp peak occurs in their plot of main-chain *B* factors at Thr211, with a maximum of more than 20 Å^2. These two factors indicate that the possibility of this peptide also being cis in the Bode structure should be investigated.

Another region of difference between the two enzyme structures is the S_4 binding site. The conformation of Tyr104 in the model of Bode et al. is close to the conformation observed in our subtilisin Novo structure (see above) and in the subtilisin BPN′ structure used as the model for their molecular replacement procedure. Tyr104 in our Carlsberg model lies closer to the enzyme surface with its phenolic ring almost perpendicular to that of Tyr104 in Novo (Figures 7 and 8). The peptide bond between Ser130 and Gly131 is reversed in Carlsberg to accommodate the changed conformation of Tyr104. The S_4 binding site in our Carlsberg model is thus adapted to fit the P_4 side chain of eglin-c, Pro56I, but retains its completely hydrophobic character (Figure 15). No additional electron density peaks for weakly bound water molecules were observed in the binding pocket, even at very low contour levels. Bode et al. see less complete enzyme–inhibitor complementarity in S_4 and report an isolated cluster of three solvent molecules at the bottom of the hydrophobic S_4 cavity. The *B*-factors for residues of the S_4 region in both Carlsberg models are in the range of 15–20 Å^2 for main-chain atoms.

Bode et al. compare their subtilisin Carlsberg structure to the subtilisin BPN′ (Wright et al., 1969) and subtilisin Novo (Drenth et al., 1972) coordinates deposited with the Brookhaven Protein Data Bank. In general, the differences and similarities between the subtilisins that they discuss are the same as those noted between our Carlsberg and Novo structures. Detailed comparisons are difficult to make because the Novo and BPN′ structures from the data bank are unrefined.

The overall shape and folding of eglin-c in the Bode structure is the same as that described in McPhalen et al. (1985a). As in the enzyme structures, the assignment of secondary structural elements in the two inhibitors (Table V) is generally the same. Bode et al. do not report on the well-ordered water molecules forming the hydrogen-bonding bridges between strand 3 and the C-terminal segment of the β-sheet in eglin-c and CI-2. In the reactive-site loop, differences of 2–9° are seen in the main-chain torsion angles ϕ and ψ between the two eglin-c structures. These values are well within the range of differences observed with other serine proteinase inhibitors (Fujinaga et al., 1987) and may be due to the use of different refinement procedures and restraints for the two complexes. The hydrogen-bonding distances between residues of the inhibitor reactive-site loops and the enzyme active sites are identical within coordinate errors. The geometry of the scissile peptide bond differs slightly in the two eglin-c models. Bode et al. report some distortion, with $\omega = 158°$ and an out-of-plane bending angle for the C–O bond (θ^4) of −10.1°. The corresponding values in our model are $\omega = 170°$ and $\theta^4 = -4.0°$, indicating no distortion within coordinate errors as discussed above.

Overall, despite differences in crystallization medium (salt vs poly(ethylene glycol) precipitation), data collection technique (diffractometer vs film), and refinement procedure [Hendrickson and Konnert (1980) vs EREF (Jack & Levitt, 1978)], the structures of the two complexes correspond closely, even down to the details of hydrogen bonding and enzyme–inhibitor contacts. It will be of interest to make a detailed comparison of these two independently determined and refined structures.

Why Are Eglin-c and CI-2 Good Inhibitors? According to the standard mechanism of action for protein inhibitors of serine proteinases put forward by Laskowski and Kato (1980), a good inhibitor binds very tightly to the enzyme (low K_M) and is hydrolyzed very slowly (low k_{cat}). The association rate constants for eglin-c with human leukocyte elastase and cathepsin G are on the order of 10^7 M^{-1} s^{-1}, and the equilibrium association constants are 10^{11} M^{-1} (Schnebli et al., 1985). The rapid and tight association of enzyme and inhibitor is promoted by the reactive-site region of the inhibitor being held fairly rigidly in a conformation complementary to the enzyme active site. The rigidity of the reactive-site region is not absolute, however, as can be seen by its high *B* factors in the structure of free CI-2 (McPhalen & James, 1987). Some loss of rigid complementarity and tightness of binding may be compensated by another characteristic of a good inhibitor, broadness of action. Moderate reactive-site flexibility allows an inhibitor to act on different enzymes by adapting its conformation to different active-site architecture.

Despite some reactive-site flexibility, both eglin-c and CI-2 seem to conform to the part of the inhibitory mechanism requiring tight binding to the enzyme. The complementarity between their reactive-site loops and the enzyme active site in the complexes is excellent. The reactive-site loop of the inhibitors is stabilized by extensive hydrogen bonding within the inhibitor. The conformations of the loops of CI-2 and eglin-c are highly similar, as are their interactions with their respective enzyme's active site. In the complex, a number of interactions favor tight binding of the inhibitor to the enzyme. These include the hydrogen bonding in the two sets of β-sheet interactions with subtilisin, as well as extensive hydrophobic interactions of inhibitor side chains with deep specificity pockets on the enzyme.

The second required property of a good inhibitor is a slow rate of hydrolysis. The K_{hyd} for inhibitors is usually close to 1 near neutral pH (Finkenstadt et al., 1974; Laskowski & Kato, 1980). Values of K_{hyd} are not available for CI-2 and eglin-c, but the similarity of their interactions in enzyme complexes to the interactions of other families of inhibitors indicates that they may share a common mechanism. The slow rate of hydrolysis of protein inhibitors of serine proteinases may be due to activation energy barriers at many points in the reaction.

Eglin-c and CI-2 share several structural features with ovomucoid inhibitors that may contribute energy barriers to hydrolysis (Fujinaga et al., 1982; Read et al., 1983). Formation of a tetrahedral intermediate may be hindered by the presence of favorable hydrogen bonds to main-chain atoms flanking the reactive-site bond (Figure 13). Barriers to relaxation of the reactive-site region after hydrolysis are provided by the stabilizing network of interactions supporting the entire reactive-site loop. The network may reduce the conformational freedom of a cleaved inhibitor, analogous to the disulfide

linkages in ovomucoid inhibitors. The charges on the new termini of a cleaved bond could be stabilized through interaction with the charged side chains of P_1' (negative) and Arg65I plus Arg67I (positive). The positive end of a helix dipole in the ovomucoid inhibitors is positioned analogously to Arg65I. An inhibitor closely related to CI-2, CI-1, has a phenylalanine residue in place of Arg67I (Svendsen et al., 1982). CI-1 is cleaved and dissociated from serine proteinases more rapidly than CI-2, indicating the importance of the arginine residues to inhibition (I. Svendsen, personal communication).

Although these common structural features among the reactive-site loops of the inhibitor families are striking, their actual significance to inhibition is uncertain. Recent kinetic data (Ardelt & Laskowski, 1985) show that the size of the overall barrier to hydrolysis may vary over several orders of magnitude for the same inhibitor with different enzymes. This indicates that the dominant individual activation energy barriers are due to enzyme–inhibitor interactions (Read & James, 1986). The more variable S′ and P′ regions may interact differently in different enzyme–inhibitor complexes. For example, the P′ segment may be more hindered by some enzymes in moving away from the active site after hydrolysis (Read & James, 1986). Information on such differential enzyme–inhibitor interactions for eglin-c or CI-2 will require three-dimensional structures of their complexes with non-subtilisin serine proteinases.

A model of eglin-c after cleavage of the reactive-site bond can be built into the active site of subtilisin Carlsberg. Starting from the coordinates of the intact inhibitor, rotations about The $N–C^{\alpha}$ bond of P_1 and the $C^{\alpha}–C$ bond of P_1' are sufficient to move the new termini to nonbonded contact distance. The carbonyl carbon of P_1 can be placed about 1.8 Å from the O^{γ} of Ser221 to stimulate the acyl enzyme. The side chain of Leu59I can be adjusted within the broad S_1 specificity pocket. Movement of the NH of P_1' away from the carbonyl carbon of P_1 shifts the side chain of P_1' away from its position in the hydrogen-bonding network, but rotations about χ_1 and χ_2 restore most of its interactions. No strong conclusions can be drawn from such a rough model, but it does indicate that the inhibitor could remain bound to the enzyme after cleavage without excessive disruption of favorable binding and stabilizing interactions.

ACKNOWLEDGMENTS

We thank H. P. Schnebli of Ciba-Geigy for the gift of purified eglin-c, I. Svendsen for the gift of purified subtilisin Carlsberg, and Koto Hayakawa for growing the crystals of the complexes. We thank Mae Wylie for cheerfully typing the manuscript throughout its many revisions. We also thank Masao Fujinaga, Randy Read, Anita Sielecki, and John Moult for many long and fruitful discussions.

REFERENCES

Ardelt, W., & Laskowski, M., Jr. (1985) *Biochemistry 24*, 5313–5320.

Barel, A. O., & Glazer, A. N. (1968) *J. Biol. Chem. 243*, 1344–1348.

Barry, C. D., Molnar, C. E., & Rosenberger, F. U. (1976) *Technical Memo No. 229*, Computer Systems Lab, Washington University, St. Louis, MO.

Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. J., Jr., Brice, M. D., Rogers, J. K., Kennard, O., Shimanouchi, T., & Tasumi, M. (1977) *J. Mol. Biol. 112*, 535–542.

Bode, W., & Schwager, P. (1975) *J. Mol. Biol. 98*, 693–717.

Bode, W., Papamakos, E., Musil, D., Seemüller, U., & Fritz, H. (1986) *EMBO J. 5*, 813–818.

Bode, W., Papamakos, E., & Musil, D. (1987) *Eur. J. Biochem. 166*, 673–692.

Brown, I. D., & Shannon, R. D. (1973) *Acta Crystallogr., Sect. A: Cryst. Phys., Diffr., Theor. Gen. Crystallogr. A29*, 266–282.

Crawford, J. L., Lipscomb, W. N., & Schellman, C. G. (1973) *Proc. Natl. Acad. Sci. U.S.A. 70*, 538–542.

Creighton, T. E. (1983) in *Proteins: Structures and Molecular Principles*, p 228, Freeman, New York.

Cromer, D. T., & Mann, J. B. (1968) *Acta Cryst, Sect. A: Cryst. Phys., Diffr., Theor. Gen. Crystallogr. A24*, 321–324.

Crowther, A. R. (1973) in *The Molecular Replacement Method* (Rossman, M. G., Ed.) International Science Review 13, pp 173–178, Gordon & Breach, New York.

Drenth, J., Hol, W. G. J., Jansonius, J. N., & Koekock, R. (1972) *Eur. J. Biochem. 26*, 177–181.

Einspahr, H., & Bugg, C. E. (1981) *Acta Crystallogr., Sect. B: Struct. Crystallogr. Cryst. Chem. B37*, 1044–1052.

Einspahr, H., & Bugg, C. E. (1984) *Met. Ions Biol. Syst. 17*, 51–97.

Finkenstadt, W. R., Hamid, M. A., Mattis, J. A., Schrode, J., Sealock, R. W., Wang, D., & Laskowski, M., Jr. (1974) *Bayer-Symp. 5*, 389–411.

Fujinaga, M., & Read, R. J. (1987) *J. Appl. Crystallogr. 20*, 517–521.

Fujinaga, M., Read, R. J., Sielecki, A., Ardelt, W., Laskowski, M., Jr., & James, M. N. G. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 4868–4872.

Fujinaga, M., Sielecki, A. R., Read, R. J., Ardelt, W., Laskowski, M., Jr., & James, M. N. G. (1987) *J. Mol. Biol. 195*, 397–418.

Hendrickson, W. A., & Konnert, J. H. (1980) in *Biomolecular Structure, Function, Conformation and Evolution* (Srinivasan, R., Ed.) Vol. I, pp 43–57, Pergamon, Oxford, U.K.

Hirono, S., Akagawa, H., Mitsui, Y., & Iitaka, Y. (1984) *J. Mol. Biol. 178*, 389–413.

Hol, W. G. J. (1971) Ph.D. Thesis, University of Gröningen.

Hol, W. G. J. (1985) *Prog. Biophys. Mol. Biol. 45*, 149–195.

Jack, A., & Levitt, M. (1978) *Acta Crystallogr., Sect. A: Cryst. Phys., Diffr., Theor. Gen. Crystallogr. A34*, 931–935.

James, M. N. G., & Sielecki, A. R. (1983) *J. Mol. Biol. 163*, 299–361.

James, M. N. G., Sielecki, A. R., Brayer, G. D., Delbaere, L. T. J., & Bauer, C.-A. (1980) *J. Mol. Biol. 144*, 43–88.

Jany, K. D., Lederer, G., & Mayer, B. (1986) *FEBS Lett. 199*, 139–144.

Johnson, C. K. (1965) ORTEP, Report ORNL-3794, Oak Ridge National Laboratory, Oak Ridge, TN.

Kabsch, W., & Sander, C. (1983) *Biopolymers 22*, 2577–2637.

Kraut, J., Robertus, J. D., Birktoft, J. J., Alden, R. A., Wilcox, P. E., & Powers, J. C. (1972) *Cold Spring Harbor Symp. Quant. Biol. 36*, 117–123.

Kretsinger, R. H. (1976) *Annu. Rev. Biochem. 45*, 239–266.

Laskowski, M., Jr., & Kato, I. (1980) *Annu. Rev. Biochem. 49*, 593–626.

Laskowski, M., Jr., Kato, I., Ardelt, W., Cook, J., Denton, A., Empie, M. W., Kohr, W. J., Park, S. J., Parks, K., Schatzley, B. L., Schoenberger, O. L., Tashiro, M., Vichot, G., Whatley, H. E., Wieczorek, A., & Wieczorek, M. (1987) *Biochemistry 26*, 202–221.

Lee, B., & Richards, F. M. (1971) *J. Mol. Biol. 55*, 379–400.

Linderstrøm-Lang, K., & Ottessen, M. (1947) *Nature (London) 159*, 807–808.

Luzzati, V. (1952) *Acta Crystallogr. 5*, 802–810.
Marquart, M., Walter, J., Deisenhofer, J., Bode, W., & Huber, R. (1983) *Acta Crystallogr., Sect. B: Struct. Sci. B39*, 480–490.
Matsubara, H., Hagihara, B., Nakai, M., Komaki, T., Yonetani, T., & Okunuki, K. (1958) *J. Biochem. (Tokyo) 45*, 251–255.
Matthews, B. W. (1987) *Biochemistry 26*, 6885–6888.
McPhalen, C. A. (1986) Doctoral Thesis, University of Alberta.
McPhalen, C. A., & James, M. N. G. (1987) *Biochemistry 26*, 261–269.
McPhalen, C. A., Schnebli, H. P., & James, M. N. G. (1985a) *FEBS Lett. 188*, 55–58.
McPhalen, C. A., Svendsen, I., Jonassen, I., & James, M. N. G. (1985b) *Proc. Natl. Acad. Sci. U.S.A. 82*, 7242–7246.
Melville, J. C., & Ryan, C. A. (1972) *J. Biol. Chem. 247*, 3445–3453.
Olaitan, S. A., DeLange, R. J., & Smith, E. L. (1968) *J. Biol. Chem. 243*, 5296–5301.
Pähler, A., Banerjee, A., Dattagupta, J. K., Fujiwara, T., Lindner, K., Pal, G. P., Suck, D., Weber, G., & Saenger, W. (1984) *EMBO J. 3*, 1311–1314.
Ramachandran, G. N., & Mitra, A. K. (1976) *J. Mol. Biol. 107*, 85–92.
Ramakrishnan, C., & Ramachandran, G. N. (1965) *Biophys. J. 5*, 909–933.
Read, R. J. (1986) *Acta Crystallogr., Sect. A: Found. Crystallogr. A42*, 140–149.
Read, R. J., & James, M. N. G. (1988) *J. Mol. Biol.* (in press).
Read, R. J., & James, M. N. G. (1986) in *Proteinase Inhibitors* (Barrett, A. J., & Salvesen, G., Eds.) pp 301–336, Elsevier, Amsterdam.
Read, R. J., Fujinaga, M., Sielecki, A. R., & James, M. N. G. (1983) *Biochemistry 22*, 4420–4433.
Read, R., Fujinaga, M., Sielecki, A., Ardelt, W., Laskowski, M., Jr., & James, M. (1984) *Acta Crystallogr., Sect. A: Found. Crystallogr. A40*, C50–C51.
Richardson, J. S. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73*, 2619–2623.
Richardson, J. S. (1981) *Adv. Protein Chem. 34*, 167–339.
Rink, H., Liersch, M., Sieber, P., & Meyer, F. (1984) *Nucleic Acids Res. 12*, 6369–6387.
Robertus, J. D., Kraut, J., Alden, R. A., & Birktoft, J. J. (1972) *Biochemistry 11*, 4293–4303.
Rossmann, M. G. (1973) *The Molecular Replacement Method*, International Science Review 13, Gordon & Breach, New York.
Russell, A. J., & Fersht, A. R. (1987) *Nature (London) 328*, 496–500.
Schecter, I., & Berger, A. (1967) *Biochem. Biophys. Res. Commun. 27*, 157–162.
Schnebli, H. P., Seemüller, U., Fritz, H., Maschler, R., Liersch, M., Bodmer, J. L., Virca, G. D., Lucey, E. C., Stone, P. G., & Snider, G. L. (1985) in *Intracellular Protein Catabolism* (Khairallah, E. A., Bond, J. S., & Bird, J. W. C., Eds.) pp 287–290, Alan Liss, New York.
Seemüller, U., Meier, M., Ohlsson, K., Müller, H.-P., & Fritz, H. (1977) *Hoppe-Seyler's Z. Physiol. Chem. 358*, 1105–1117.
Seemüller, U., Eulitz, M., Fritz, H., & Strobl, A. (1980) *Hoppe-Seyler's Z. Physiol. Chem. 361*, 1841–1846.
Seemüller, U., Dodt, J., Fink, E., & Fritz, H. (1986) in *Proteinase Inhibitors* (Barrett, A. J., & Salvesen, G., Eds.) pp 347–355, Elsevier, Amsterdam.
Shrake, A., & Rupley, J. A. (1973) *J. Mol. Biol. 79*, 351–372.
Sielecki, A. R., Hendrickson, W. A., Broughton, C. G., Delbaere, L. T. J., Brayer, G. D., & James, M. N. G. (1979) *J. Mol. Biol. 134*, 781–804.
Sielecki, A. R., James, M. N. G., & Broughton, C. G. (1982) in *Crystallographic Computing* (Sayre, D., Ed.) pp 409–419, Clarendon, Oxford, U.K.
Smith, E. L., DeLange, R. J., Evans, H., Landon, M., & Markland, F. S. (1968) *J. Biol. Chem. 243*, 2184–2191.
Svendsen, I., Martin, B., & Jonassen, I. (1980) *Carlsberg Res. Commun. 45*, 79–85.
Svendsen, I., Boisen, S., & Hejgaard, J. (1982) *Carlsberg Res. Commun. 47*, 45–53.
Wells, J. A., Ferrari, E., Henner, D. J., Estell, D. A., & Chen, E. Y. (1983) *Nucleic Acids Res. 11*, 7911–7925.
Wright, C. S. (1972) *J. Mol. Biol. 67*, 151–163.
Wright, C. S., Alden, R. A., & Kraut, J. (1969) *Nature (London) 221*, 235–242.